On stochastic models for the spread of infections

Trapman, J.P.

2006

document version
Publisher's PDF, also known as Version of record

On stochastic models for the spread of infections

Jan Pieter Trapman

ISBN-10: 90-9020910-7
ISBN-13: 978-90-9020910-4
printed by PrintPartners Ipskamp, Enschede
Engraving on cover: La peste di Firenze nel 1348

Thomas Stieltjes Institute
for Mathematics

VRIJE UNIVERSITEIT

On stochastic models for the spread of infections

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op woensdag 13 september 2006 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Jan Pieter Trapman

geboren te Hardinxveld-Giessendam

promotoren: prof.dr. R.W.J. Meester
prof.dr. J.A.P. Heesterbeek

I returned, and saw under the sun, that the race is not to the swift, nor the battle to the strong, neither yet bread to the wise, nor yet riches to men of understanding, nor yet favour to men of skill; but time and chance happeneth to them all.
(Ecclesiastes 9:11, KJV)

Contents

Chapter 1

General introduction

1.1 Why mathematical epidemiology?

Recent outbreaks of animal diseases like classical swine fever (1997), foot and mouth disease (2001) and avian influenza (2003) in The Netherlands, the world-wide emerging of human diseases like SARS in 2002 and the constant threat of new pandemic infections, like avian influenza, make it needless to stress the importance of understanding and predicting the dynamics of the spread of an infection. Mathematical models and their analysis play a natural role in obtaining such understanding.

Computer-intensive models can be used for predictions, but many of those models are mathematically intractable. Therefore, a challenge to epidemiologists, modellers and mathematicians is to find models that are both useful and mathematically tractable.

The body of epidemiological literature is already immense. Although most of the models and methods cannot immediately be applied to epidemics in real-life, they do provide tools for the analysis of these real epidemics and give qualitative insight in the dynamics of the spread of the infection. In addition, real-life applications call for more advanced mathematical methods to be developed.

Apart from the modelling of the spread of infections, results from epidemiology have been used in other sciences as well. One can think of models for rumour spread [26], the spread of information in economy [37], the spread of computer viruses [17, 31] and even an application of an epidemic model in a model on the spread of heresies in the Middle Ages [65].

In this thesis I focus on the main purpose of mathematical epidemiology: modelling the spread of infections. Notably, I develop new methods to provide a better foundation for this modelling. The following sections of this general introduction are used for introducing some basic deterministic and stochastic epidemic theory and some theory of branching processes. Important epidemiological terms and quantities (like the basic reproduction ratio R_0) are introduced here. Readers familiar with the material can skip this part of the thesis without problems. After that a short introduction to the theory of percolation is given. This field in probability theory is mainly used in mathematical physics, but in this thesis an application of the theory in epidemiology is given. This chapter ends with an overview of this thesis and a section on possible directions for further development of (stochastic) mathematical epidemiology.

1.2 Deterministic epidemic models

In 1927, 1932 and 1933 W.O. Kermack and A.G. McKendrick published a series of papers titled "Contributions to the mathematical theory of epidemics". Those papers are often seen as the basis of further research in mathematical (especially deterministic) modelling of the spread of infectious diseases. The first three papers of Kermack and McKendrick are reprinted in [46]. In the book of Diekmann and Heesterbeek [27] some of the results of Kermack and McKendrick and other deterministic models are presented and explained.

In many of the deterministic models randomly mixing populations are assumed, i.e. every individual contacts every other individual with the same rate, or in a multi-type process, contacts between individuals of certain types all happen at the same rate.

First, I consider the basic single-type SIR (Susceptible $\rightarrow$ Infectious $\rightarrow$ Removed) epidemic model. Generalisation to multi-type models is often straightforward, but the notation becomes more cumbersome. In the model the population is closed and is of size $\mathcal{N}$, i.e. births, deaths and migration are ignored. All pairs of individuals have contacts at the same rate $\mathcal{N}^{-1}\beta$. If an infectious individual has a contact with a susceptible individual, the susceptible individual becomes infectious. Infectious individuals recover at rate α.

Define $S(t)$ as the number of individuals susceptible at time t and $s(t) = \mathcal{N}^{-1}S(t)$ as the fraction of the population that is susceptible at time t. $I(t)$, $i(t)$, $R(t)$ and $r(t)$ are defined similarly. Let $i(0) = i_0$ and $s(0) = 1 - i_0$.

This leads for large $\mathcal{N}$ to the following set of approximating differential equations:

$$\begin{aligned}\frac{ds(t)}{dt} &= -\beta s(t)i(t),\\ \frac{di(t)}{dt} &= \beta s(t)i(t) - \alpha i(t) = i(t)(\beta s(t) - \alpha),\\ \frac{dr(t)}{dt} &= \alpha i(t).\end{aligned}$$

This is the place to introduce the concept of the basic reproduction ratio (or basic reproduction number) R_0. The basic reproduction ratio is defined as the number of secondary infections per initial infective individual, in a very large population where $s(0) \approx 1$. In this model an initially infected individual will on average infect

$$\frac{(\mathcal{N}-1)\beta}{\mathcal{N}\alpha} \approx \frac{\beta}{\alpha}$$

other individuals. So $R_0 \approx \beta/\alpha$.

Some remarks have to be made:

- It can be seen from the differential equations that if a small number (relative to the total population size) of infective individuals is introduced in a completely susceptible population, the number of infectives will initially grow if $\beta - \alpha > 0$, i.e. if $R_0 \approx \frac{\beta}{\alpha} > 1$. On the other hand, if $R_0 \leq 1$ the epidemic will go extinct very soon. This is the reason why R_0 is such an important quantity in epidemiology: It is a threshold parameter, like the offspring mean in branching processes (see Section 1.4).
- It is proven [27] that $\log(s(\infty)) = R_0(s(\infty) - 1)$. This does not lead to a closed expression for $s(\infty)$, but it can be computed numerically. Note that $s(\infty)$ is the fraction of the total population that is never infected.
- Other deterministic models for infection spread usually lead to a comparable set of differential equations. In particular, it is easy to incorporate birth and death in the model.
- If birth and death are considered in the model we can also deduce conditions for an epidemic to become endemic, i.e. an epidemic for which $i(t) \nrightarrow 0$ as $t \to \infty$.
- A varying environment can be treated by making β and α time-dependent.

- The duration of an epidemic cannot be studied in a realistic way with a deterministic model, because at the start and at the end of an epidemic stochastic fluctuations play a very important role.
- For some infections it is not enough to use three states of infection (susceptible, infective and removed), but we also need to take a latent state or more than one infectious states into account. One can deal with this by using a system of more differential equations than the three already given, but the main idea is still the same.

The notion of the basic reproduction ratio R_0 is very important for controlling an epidemic. By vaccinating a fraction larger than $1 - \frac{1}{R_0}$ of a well mixed population at birth with a perfect vaccine, one can bring the reproduction ratio in the vaccinated population below 1 and thus prevent a major outbreak.

An interesting problem is the modelling of spread of an infection among households. In this case two levels of infection are considered, the probability of infecting others within a household is higher than the probability of infecting individuals outside the own household. Note that for small households it is not very realistic to use a purely deterministic model. In this setting it is also very difficult to take a varying environment into account. The household model is also used in stochastic models [6], but in those models it is even more difficult to deal with varying environments (see Chapter 2).

1.3 Stochastic epidemic models

Deterministic models have drawbacks and may cause some rather amusing mistakes: Murray et al. ([60], see also page 693 of [61]) proposed a deterministic model to predicted the dynamics of the prevalence of rabies among foxes in England. The dynamics predicted by the model seems to be rather spectacular. The number of infected foxes will rapidly increase until the number of available susceptible foxes is too low and then the disease seems to disappear, but after a period of about 2 years there is a sudden reappearing of rabies in foxes predicted. In 1991 Mollison [59] gave an explanation for this phenomenon, which he called the atto-fox phenomenon. Murray et al. have used a continuous approximation of the number of infected foxes (like is done in the differential equations in the deterministic SIR-model) and during the years

that the infection seemed to have disappeared (in the predictions based on the model), there still was a minimum of around one infected atto-fox (10^{-18} of a fox) per square kilometre and this fox-part eventually caused the new wave of infection.

This example shows the danger of continuous approximations of the number of individuals. Furthermore, as explained in the previous section for discrete models one has to assume that the number of infectious individuals is large and that the number of possible contacts with susceptible individuals is large as well. Therefore, deterministic models are not the most relevant for modelling the start and the end of an epidemic (where the number of infectious individuals is small) and for modelling the spread of infections on networks (where the number of available susceptible individuals that an infective individual can infect, is small). In this thesis we will mainly focus on these problematic cases i.e. we focus on stochastic and discrete models.

Much of the work already done on stochastic epidemic models is described and explained in the lecture notes of Andersson and Britton [3]. In this section I give some of the results and the methods they describe.

1.3.1 The Reed-Frost model

The most basic model is the "Reed-Frost model". This model is generation-based. We assume that the population is closed, i.e. during the epidemic no new individuals are introduced in the population and there are no deaths. The probability of a susceptible to become infectious in generation $i+1$ only depends on the number of infectives in generation i. We assume that the probability of escaping infection by a specific infective is q for each susceptible. Conditioned on the number of infectives in generation i the events that different susceptibles escape from the infection in generation $i+1$ are independent. Note that in the last assumption, a fixed infective period and the same infectivity rate for all infectives is implicitly assumed. We get:

$$\mathbb{P}(I(j+1)=n|S(j)=k,I(j)=l)=\binom{k}{n}(1-q^l)^n(q^l)^{k-n}$$

(and $S(j+1)-S(j)=I(j+1)$).

We are interested in the final size (the number of initially susceptible individuals that are ultimately removed) of an epidemic. We can write this as:

$$\mathbb{P}(R(\infty) = k|S(0) = n, I(0) = m) = \\ \sum_{l:|l|=k} \mathbb{P}(I(1) = l_1, \ldots, I(j) = l_j, I(j+1) = 0|S(0) = n, I(0) = m),$$

where $|l| = \sum_{j\geq 1} l_j$. For practical purposes this formula is useless for large populations, because the needed number of computations grows very fast and worse: they are numerically unstable.

1.3.2 The standard stochastic SIR epidemic model

For the stochastic SIR model we have the following assumptions:

- The population is closed, so no births, deaths and migration are considered in the model.
- Initially there are $\mathcal{N} - m$ susceptibles and m infective individuals.
- The population is homogeneous and randomly mixing, i.e. the probability that two individuals contact each other does not depend on which two individuals we consider and all individuals have the same characteristics.
- At each contact between an infective and a susceptible the infection is transmitted.
- Each individual contacts a given other individual at the time points of a Poisson process with parameter $\frac{\beta}{\mathcal{N}}$. With some abuse of terminology β is called the infection rate or contact rate.
- An individual is in the R-class if its infectious period is terminated.
- Individuals have independent and identically distributed (i.i.d.) infectious periods, all distributed as the random variable $\mathcal{I}$ with mean ι and finite variance σ^2.

The process described above is denoted by $E_{\mathcal{N}-m,m}(\beta, \mathcal{I})$.

One remark on the terminology: from now on in this thesis the word "rate" in a stochastic setting will mean the density of a one dimensional Poisson process.

The model proposed in this section slightly differs from models proposed in the literature (e.g. [3]). Andersson and Britton and many other authors

working on stochastic epidemics assume that the contacts are made according to a Poisson-process with parameter $\frac{\beta}{\mathcal{N}-m}$. In this model a process started with 2 infective and n susceptible individuals will behave in another way than a process initiated by 1 infective, and $n+1$ susceptibles, given that one of the initially susceptible individuals is infected immediately. Note that replacing β by $\beta(\mathcal{N}-m)/\mathcal{N}$ in the model proposed in [3] brings us back at the standard stochastic SIR epidemic model described above.

The basic reproduction number, R_0, is given by $\beta\iota$, i.e. the rate at which an infective individual contacts susceptible ones times the expected length of the infective period, or better: The expected number of secondary infected individuals directly infected by a single initially infected individual.

We can look at the SIR model from a different perspective. Sellke gives such an alternative construction of the model [68]. In Sellke's construction each susceptible is associated with a threshold. As soon as the total infective pressure (i.e. $\frac{\beta}{\mathcal{N}}\int_0^t I(u)du$: the number of infective individuals times the infection rate integrated up to time t) is above the threshold the susceptible becomes infective and adds to the total infectivity pressure. More formally: The initial infectives are labelled: $-(m-1), -(m-2), \ldots, 0$ and the initial susceptibles are labelled: $1, 2, \ldots, n$, where $n = \mathcal{N} - m$. We introduce random variables for the (possible) infectious periods, $\mathcal{I}_j$ for $-(m-1) \leq j \leq n$, all distributed as $\mathcal{I}$. The initially susceptibles each get an exponential distributed random variable with mean one, Q_j associated to it. The number of infectious individuals at time t is $I(t)$. Let

$$A(t) := \frac{\beta}{\mathcal{N}} \int_0^t I(u)du$$

be the total infection pressure exerted on a given susceptible up to time t. Now the susceptible labeled j becomes infective as soon as $A(t)$ is greater or equal than Q_j. The j-th individual that becomes infected (not necessarily the susceptible labeled j) stays infective during a period $\mathcal{I}_j$, after this period it is removed. As soon as there are no more infectives the epidemic ceases. It is proved [3] that this model is equivalent to the SIR-model constructed earlier.

Note that we can use this construction for the final size equation:

$$Z := R(\infty) - m = \min\Big\{j | Q_{(j+1)} > \frac{\beta}{\mathcal{N}} \sum_{k=-(m-1)}^{j} \mathcal{I}_k\Big\}.$$

Here $Q_{(j)}$ denotes the j-th order statistic of the Q_j's. The total infection pressure up to the end of the epidemic is now given by

$$A := A(\infty) = \frac{\beta}{\mathcal{N}} \sum_{k=-(m-1)}^{Z} \mathcal{I}_k. \tag{1.1}$$

Andersson and Britton give the following results (first given by Ball in [4]):

Lemma 1.3.1 *Consider the standard SIR epidemic $E_{n,m}(\beta, \mathcal{I})$ and let A be the total infection pressure of the epidemic. Then*

$$\mathbb{E}[e^{-\theta A}/\phi(\beta\theta/\mathcal{N})^{Z+m}] = 1, \qquad \theta \geq 0,$$

where $\phi(\theta) = \mathbb{E}(e^{-\theta\mathcal{I}})$ is the Laplace transform of $\mathcal{I}$.

This lemma, which is interesting on its own, is used to prove the following result:

Theorem 1.3.2 *Consider the standard SIR epidemic $E_{n,m}(\beta, \mathcal{I})$. Denote by P_k^n the probability that the final size of the epidemic is equal to k, $0 \leq k \leq n$. Then*

$$\sum_{k=0}^{l} \binom{n-k}{l-k} P_k^n / [\phi(\beta(n-l)/\mathcal{N})]^{k+m} = \binom{n}{l}, \qquad 0 \leq l \leq n.$$

With this theorem we can find the final size probabilities recursively. Note that we do not need to assume the absence of a latent period and we may replace β by $\beta(\tau)$, where τ is the time since infection of an individual.

1.4 Branching processes

Two of the chapters (2 and 3) of this thesis have branching processes as their main subject, while in Chapter 5 basic theory from branching processes is heavily used to analyse infection spread on random networks. Indeed, there is a strong relation between epidemiology and the theory of branching processes, or to do more justice to the facts: mathematical epidemiology heavily relies on the theory of branching processes and draws inspiration from it. Branching processes were first developed to describe the extinction probabilities of family

names and can be seen as the mathematical study of family trees. (For an overview of the history of branching processes see the introduction of [39].)

The basic (Galton-Watson) branching process model can roughly be seen as follows: There are some (male) ancestors that each have a random number of sons. The number of sons of the different ancestors are independently and identically distributed (i.i.d.). The sons of the ancestors have sons themselves, and the numbers of these sons are again i.i.d. and distributed as the offspring of an ancestor. In this way we can define a model for the whole family tree. All individuals have independently a number of sons, all according to the same distribution.

The model can be enriched by considering the family tree in real time. In such an extended model not only the generation number of an individual is important, but also the times of birth and death are. Such a model is useful (and even needed) if, for example, one wants to make predictions about the number of people with surname Johnson in the year 2100 (Of course under the assumption that life expectancy, the distribution of the number of sons and the age at which people get their sons does not change between now and 2100).

The relation between epidemics in large randomly mixing populations and family trees can heuristically be explained as follows. The individuals that introduce an infection into a population are seen as the ancestors. During their infectious period they contact a random number of individuals uniformly chosen from the population and infect those individuals if they are still susceptible. These newly infected individuals can be seen as generation 1 individuals. During their infectious period, these generation 1 individuals contact the other individuals according to the same law as the ancestors contacted other individuals.

If the population is very large, the probability that an infective individual contacts an already infected individual during the first stage of the epidemic (which may last rather long) is very small, and therefore the progress of the epidemic can be described by a branching process.

In the next subsection the so-called Galton-Watson processes are rigorously defined and some of the basic results on these processes are given. Furthermore, some extensions, like multi-type Galton-Watson processes, are defined and generalisations of the results on Galton-Watson processes are given. For proofs of the results and for further theory on branching processes one might

consult [39]. In subsection 1.4.6 a theorem from [5] on the approximation of epidemics by branching processes is given.

From now on we use the common convention of using the maternal terminology, so we speak of mothers that give birth to daughters.

1.4.1 Definition of the Galton-Watson process

In this subsection we consider branching processes seen from a generation point of view, namely Galton-Watson processes. We assume that there is only one ancestor. This assumption is just for notational convenience and can easily be dropped.

A Galton-Watson process is a process generating a family tree, where every individual has a finite (not necessarily bounded) number of daughters. The ancestor is labelled by a. The j_n-th daughter of the j_{n-1}-th daughter $\dots$ of the j_1-th daughter of the ancestor is labelled by the vector $(a, j_1, \dots, j_n)$. If an individual can give birth to more than one daughter at once, one can use a random ordering to label the daughters.

The set of all possible labels J is given by

$$J = \{a \cup \bigcup_{n=1}^{\infty} \{(a; x); x \in \mathbb{N}^n\}\},$$

where $\mathbb{N}$ is the set of positive integers (0 not included). This set of labels is countable.

From now on "individual x" will mean the individual with label $(a, x) \in J$ and "individual a" is the ancestor.

Note that there are far too many labels, indeed if the ancestor has only 1 daughter, there are no individuals with the labels of the form

$$(a, 2, j_2, \cdots j_n; n \in \mathbb{N}, j_i \in \mathbb{N}).$$

Let r_x be 1 if the individual x is realised, i.e. $r_x = 1$ if the label (a, x) is assigned to an individual. Otherwise $r_x = 0$.

Let $\{\xi_x, x \in J\}$ be a set of random variables that are i.i.d. and distributed as a random variable ξ. If $r_x = 1$, then ξ_x denotes the total number of daughters of individual x. Furthermore, let $p_k := \mathbb{P}(\xi = k)$.

Now define the random variables X_n for $n \in \mathbb{N}$,

$$X_n := \sum_{x \in \mathbb{N}^n} r_x = \sum_{x \in \mathbb{N}^{n-1}} r_x \xi_x,$$

as the total number of individuals in generation n. Let X_0 be the number of ancestors, which we have assumed to be 1. $\{X_n; n \in 0 \cup \mathbb{N}\}$ form a homogeneous Markov chain.

1.4.2 The generating function and the extinction probability

The (probability) generating function of a discrete random variable R [32, 39] is defined as

$$f_R(s) := \mathbb{E}(s^R) = \sum_{i=0}^{\infty} \mathbb{P}(R = i) s^i$$

for all $s \in \mathbb{R}_{\geq 0}$ for which this sum converges. If $s \leq 1$ the sum converges. The generating function has useful properties like

$$\begin{aligned} f'_R(1) &= \mathbb{E}(R), \\ f''_R(1) &= \mathbb{E}(R(R-1)), \\ f_R^{(k)}(0) &= k! \mathbb{P}(R = k), \end{aligned}$$

where $f_R^{(k)}(s)$ is the k-th derivative of $f_R(s)$ and $f'_R(1) = \lim_{s \uparrow 1} f'_R(s)$ if the generating function is not defined for $s > 1$. The last property shows that $f_R(s)$ determines the distribution of R.

For the analysis of Galton-Watson processes we consider the following generating functions

$$\begin{aligned} f(s) &:= f_\xi(s) = \mathbb{E}(s^\xi), \\ f_k(s) &:= f_{X_k}(s) = \mathbb{E}(s^{X_k}) \end{aligned}$$

for $k \in \mathbb{N}$.

Note that

$$\begin{aligned} f_{k+1}(s) &= \mathbb{E}(s^{X_{k+1}}) \\ &= \mathbb{E}(\mathbb{E}(s^{X_{k+1}}|X_k)) \\ &= \mathbb{E}([\mathbb{E}(s^{\xi})]^{X_k}) \\ &= \mathbb{E}([f(s)]^{X_k}) \\ &= f_k(f(s)) \end{aligned}$$

and by induction we obtain $f_{k+1}(s) = f(f_k(s))$.

The generating function is extremely important in the theory of Galton-Watson processes, because it is the key to finding the extinction probability of the process, and it is the quest for extinction probabilities that led to the development of the theory of branching processes.

Some simple arguments can explain the relation between the generating function and the extinction probability. Let q be the probability that the progeny of a single ancestor goes extinct. If the progeny of the ancestor goes extinct this means that she has no daughters at all or that the progeny of all of her daughters goes extinct. However, the number of daughters of different individuals are i.i.d. This brings us to the equation

$$q = \sum_{i=0}^{\infty} p_i q^i = \mathbb{E}(q^{\xi}) = f(q). \tag{1.2}$$

This argument has been used in Chapter 2 to obtain equations to compute the probability of a major outbreak of an epidemic.

If the offspring mean $m := \mathbb{E}(\xi)$ is larger than 1, then the equation above has two roots in $[0, 1]$. In that case, it is well known [39] that the extinction probability is the smallest of the two. The other root is always 1. If $m \leq 1$ and $p_1 \neq 1$, then 1 is the only root in $[0, 1]$. Galton-Watson processes with $m = 1$ are called critical processes, and if m is greater or less than 1 the processes are respectively supercritical and subcritical.

An important property of Galton-Watson processes is that with probability 1 they either go extinct or explode, i.e. with probability 1, $\lim_{k\to\infty} X_k$ is either 0 or ∞ (Theorem 2.3.2 of [39]).

In this thesis I mainly consider supercritical branching processes. In Chapter 2 the extinction probability of a special branching process is computed. In

Chapter 3 moments of the offspring distribution are estimated. The estimators only converge on the set where the process does not go extinct. In Chapter 5 we use branching processes to determine features like the reproduction ratio R_* (a quantity similar to R_0, which in turn can be interpreted as the offspring mean m) and the extinction probability for the spread of an infection on random networks.

1.4.3 Martingale convergence

Martingales play an important role in the theory of branching processes, especially for proving convergence results one can hardly do without. This becomes apparent in Chapter 3 where we make frequent use of martingales in order to prove convergence of estimators for the moments of the offspring distribution. The theorems used is Chapter 3 indirectly use the following basic theorem from Branching processes (compare this with Theorems 2.7.1 and 2.7.3 of [39]),

Theorem 1.4.1 *If $\mathbb{E}(\xi \log(\xi)) < \infty$ the martingale $W_n := m^{-n} X_n$ converges almost surely and in L^1 norm to a non-degenerate random variable W, which is 0 if and only if $X_n \to 0$.*

1.4.4 Multi-type Galton-Watson processes

In this subsection some results from the theory of multi-type Galton-Watson branching processes are given (See Chapter 4 of [39]). We consider processes with only one ancestor. Results can easily be generalised to processes with more ancestors.

There exist results similar to those of single-type branching processes for r-type Galton-Watson processes. Let $X_n(i)$ be the vector with the number of individuals of the different types in generation n, given that the ancestor is of type i. For $1 \leq i, j \leq r$, let m_{ij} be the mean number of j-daughters of one i-individual (i.e. $m_{ij} = \mathbb{E}[(X_1(i))_j]$). This defines an $r \times r$ matrix m. We assume that the matrix m is positively regular, i.e. there exists an $n > 0$ such that all entries in m^n are strictly positive. Let ρ be the positive eigenvalue of m which is greater than or equal to all other eigenvalues in absolute value (this eigenvalue exists by the Perron-Frobenius theorem for irreducible matrices). Furthermore we use the following vector notation.

$$\begin{aligned} s &= [s_1, \ldots, s_r] \in [0,1]^r, \\ \tilde{k} &= [k_1, \ldots k_r] \in \mathbb{Z}_+^r, \\ \tilde{0} &= [0, \ldots, 0] \in \mathbb{Z}_+^r, \\ f^{(i)}(s) &= \sum_{\tilde{k}} \mathbb{P}((X_1(i))_j = k_j; 1 \leq j \leq r) s_1^{k_1} \ldots s_r^{k_r} \qquad \text{for } 1 \leq i \leq r, \\ f(s) &= [f^{(1)}(s), \ldots, f^{(r)}(s)]. \end{aligned}$$

Let q_i be the probability of extinction of the branching process given that it started with one i individual. From Theorem 4.2.2 of [39] we know that the $[q_1, \ldots, q_r]$ is the solution of $f(s) = s$ with the smallest Euclidean distance to the origin. If $\rho > 1$, then $q_i < 1$ for all i and if $\rho \leq 1$, then $q_i = 1$ for all i.

Furthermore, Theorem 4.2.6 of [39] states that if $\rho > 1$ then $\rho^{-n} X_n(k) \to W_k v$, where v is the left eigenvector of m and W_k is a random variable with the property that $\{W_k > 0\}$ differs from $\{\lim_{n\to\infty} X_n(k) \neq \tilde{0}\}$ only on a null-set.

1.4.5 General branching processes

Galton-Watson processes are used to describe a family tree from the perspective of generations. However in epidemiology one may be more interested in the real time development of the number of infective individuals. Usually it is not possible to observe the size of an "infection generation", but it may be possible to observe the number of infective individuals at a certain time.

The theory of branching processes goes further than only Galton-Watson processes, and it is possible to define a branching process in real time. However, results are much harder to obtain. The process is defined as follows. Assign i.i.d. a pair (λ_x, η_x) to every label $x \in J$. If x is realised in the branching process, the non negative random variable λ_x is the life length of individual x, while η_x is a point process on the non-negative real line denoting the reproduction of individual x. Let $\eta_x([t_1, t_2])$ denote the number of points of η_x in the interval $[t_1, t_2]$ and use $\eta_x(t)$ as a shorthand for $\eta_x([0, t])$. The total number of daughters of x is given by $\eta_x(\lambda_x) = \eta_x(\infty)$. From now on, we drop the subscript x, if it does not lead to confusion. We define $\mu([t_1, t_2]) := \mathbb{E}(\eta([t_1, t_2]))$ and $\mu(t) := \mathbb{E}(\eta(t)) < 1$. For technical reasons we assume that the process is non-lattice i.e. there is no $d > 0$ such that $\sum_{k=0}^{\infty} \mu(\{kd\}) = \mu(\infty)$. Furthermore, we assume that $\mu(\infty) < \infty$ and $\mu(0) < 1$. It is proven [39] that this

last assumption is enough to guarantee that at any given time the number of living individuals is finite.

We are interested in the number of living individuals younger than a at time t, X_t^a. The following theorem from [39] can be used:

Theorem 1.4.2 *For a general, non-lattice branching process with $\mu(\infty) > 1$ and $Var[\eta(\infty)] < \infty$, let α be the Malthusian parameter, i.e. the root of*

$$\int_0^\infty e^{-\alpha t}\mu(dt) = 1.$$

It holds that $\lim_{t\to\infty} e^{-\alpha t}X_t^a$ exists a.s. and in mean square. The limit can be written as

$$\lim_{t\to\infty} e^{-\alpha t}X_t^a = Z\int_0^a e^{-\alpha t}\mathbb{P}[\lambda > t]dt/\beta = Z\{1 - \mathbb{E}[e^{-\alpha(a\wedge t)}]\}/(\alpha\beta),$$

where $\beta = \int_0^\infty te^{-\alpha t}\mu(\{dt\})$ is the mothers average age of child bearing and Z is a random variable such that $\mathbb{E}[Z] = 1$ and $\mathbb{P}[Z = 0, X_t^\infty \nrightarrow 0] = 0$.

1.4.6 Approximating epidemics by branching processes

The start of an SIR epidemic in a large population can be approximated by the start of a continuous time branching process. Ball and Donnelly give some useful results in [5]. I give one of those results here:

Theorem 1.4.3 *There is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined a sequence of epidemic models indexed by n (the initial number of susceptibles) and the approximating branching process, with the following properties:*
Let $I_n(t)$ be the number of infectives at time t in a population of size n. And $X(t)$ the number of individuals in the branching process at time t.
Denote by A the set on which the branching process $X(\cdot)$ does not go extinct:

$$A = \Big\{\omega \in \Omega : \lim_{t\to\infty} X(t,\omega) \neq 0\Big\}.$$

We use A^c to denote the complement of A. Then, as $n \to \infty$,

$$\sup_{0\le t\le\infty} |I_n(t) - X(t)| \to 0$$

for $\mathbb{P}-$almost all $\omega \in A^c$.

Further for any $c_1 < (2\alpha)^{-1}$ *and* $c_2 > (2\alpha)^{-1}$,

$$\sup_{0 \le t \le c_1 \log n} |I_n(t) - X(t)| \to 0$$

and

$$\sup_{0 \le t \le c_2 \log n} |I_n(t) - X(t)| \to \infty$$

as $n \to \infty$, *for* $\mathbb{P}-$*almost all* $\omega \in A$. *Here* α *is the Malthusian parameter and* $\mu(t)$ *is the expected number of "children" of an ancestor of the branching process up to time* t *as before.*

This theorem implies that if the initial number of susceptible individuals goes to infinity, then the probability of a small outbreak in the population converge to the probability of extinction in the corresponding branching process. Furthermore, if there is a small outbreak, then the distribution of the final size of the epidemic is in the limit equal to the distribution of the final size of the corresponding branching process on the extinction set. Finally it is shown that epidemic processes are well described by branching processes up to a time of order $\log n$.

1.5 Percolation

Consider a very large orchard with trees planted at regular distances in such a way that the positions of the trees can be seen as the vertices of the square lattice. Now assume that one of the trees somewhere near the centre of the orchard is infected by a disease. The infection has the following characteristics. Exactly one time unit after being infected a tree will die. As long as a tree is infectious (i.e. until its death) it will spread infectious material to its nearest neighbours, which thereby become infected if they were not already so. The probability that a given nearest neighbour receives infectious material from the infectious tree is p and infections of nearest neighbours happen independently of each other.

No infectious material is spread farther than the nearest neighbours. Note that the infection under consideration has SIR dynamics, where the R state stands for death. To model the spread of such an infection on a square lattice we use percolation theory. (For a rigorous and extensive treatment of the subject see [33]). In the percolation model the dynamic character of the

epidemic spread has been dropped and only the static clusters of the spread are considered.

1.5.1 The model

The "network" that is mainly studied in percolation theory is the graph where the vertices (sites, nodes) are the points of the regular lattice $\mathbb{Z}^d$ and edges (bonds, connections) are drawn between any two vertices at Euclidean distance 1. However, other networks can be studied as well. Vertices connected by an edge are neighbours. *Edges* are open with probability p independently of each other. If an edge is not open then it is closed. Examples of question to be answered are: Is there a positive probability that one can reach infinitely many vertices from the origin by crossing only open edges? (i.e. is the origin contained in an infinite open cluster?) How does this probability depend on p? Does there exists a value p_c, strictly between 0 and 1, with the property that the probability the origin is part of an infinite open cluster is 0 for $p < p_c$ and strictly larger than 0 if $p > p_c$?

The model described above is called *bond percolation*. We can also consider *site percolation*, where *vertices* are open with probability p independently of each other. The open cluster of the origin here consists of the vertices that can be reached by paths consisting of edges with two open end-vertices.

In more mathematical notation we use the measure $\mathbb{P}_p$ for the measure associated to the bond percolation model described above. We drop the subscript if we are not immediately interested in how the probability of an event dependent on p. The probability that the origin is part of an infinite cluster is denoted by $\theta(p) := \mathbb{P}_p(0 \leftrightarrow \infty)$ and $p_c := \inf\{p : \theta(p) > 0\}$. Kesten [45] proved that for bond percolation on $\mathbb{Z}^2$ the critical probability is given by $p_c = 1/2$. For site percolation p_c on the square lattice is unknown.

In fact we should have mentioned the underlying graph in the above notation as well, but in general it is clear which graph we consider. Note that we are using product measure, so generalisations to other graphs are straightforward.

For epidemiological purposes it is more natural to consider a directed underlying graph. In such a graph edges are replaced by directed edges. The presence of an edge from vertex v_1 to vertex v_2 does not necessarily imply that there is an edge from v_2 to v_1, although in some models (like the one presented in Chapter 5) only graphs where this implication holds, are considered. It is

still the product measure that is considered, so whether or not the edge from v_1 to v_2 is open is independent of the "state" of the other edges, in particular it is independent of the state of the edge from v_2 to v_1. The cluster of the origin consists of those vertices that can be reached by an open path from the origin.

1.5.2 A relation between percolation and SIR epidemics

In this subsection we consider undirected graphs. One can explore the static (bond) percolation cluster of the origin in the following dynamic way. The construction of a percolation cluster on directed graphs is similar to the construction below. Say that the origin is the *generation 0 vertex*. Now explore the states (open or closed) of all the edges with the origin as one of the endpoints. The other endpoints of the explored open edges are the generation 1 vertices. The following step is to explore all the edges with a generation 1 vertex as endpoint, that were not explored before. The endpoints of the newly explored open edges are the generation 2 vertices, if there was no other generation number assigned to it yet. If a vertex has a generation number assigned to it, it keeps that number for ever. The $n+1$-th step is to explore the states of all the edges with a generation n vertex as end vertex, that were not explored before, and say that all the end vertices of the open edges that had not assigned a generation number to it yet, are generation $n+1$ vertices. This procedure will stop in a finite number of steps if and only if the cluster of the origin is finite.

Now return to the example of the orchard above and explore whether or not infectious material is transmitted between a pair of trees. First explore the trees next to the initial infective tree. If infectious material is transmitted to such a tree from the initial infective tree, it is said to be a generation 1 tree. Now explore whether or not infectious material is transmitted from the generation 1 trees to its neighbours for the pairs that have not been explored before. In such a way the whole cluster of infected trees can be explored, and in fact this exploration procedure is exactly the same as the exploration procedure of the percolation cluster. So the probability of an infinite outbreak of the tree-infection in an infinite orchard is $\theta(p)$. In particular, if the probability that a tree transmits infectious material to a given neighbouring tree is larger than 0.5, there is a positive probability of a large outbreak of the infection.

Here we use that it does not matter for the size of the cluster whether or not the edges are directed. Because the second endpoint of an edge of which the corresponding edge in opposite direction has already been explored is not susceptible anymore.

The exploration procedure can easily be generalised to more than one (but finitely many) initial infectives on more general graphs that are locally finite (i.e. graphs where each vertex has a finite number of neighbours).

1.5.3 Locally dependent percolation

The main extension of percolation theory that is used in this thesis, is "locally dependent percolation" by which locally dependent random graphs are constructed (see Chapter 5 and [25, 51]). We consider percolation on a directed graph, where each vertex has a finite number of edges starting at it. We introduce randomness at two levels. First we assign "infectious periods" $\mathcal{I}(v)$ independently and identically distributed to each vertex. The second step is declaring the edges in the graph "open" or "closed" again using product measure, *but* in this model the edge from v_1 to v_2 is open with probability $p(\mathcal{I}(v_1)) = 1 - exp[-\tau\mathcal{I}(v_1)]$, for some "infection rate" τ. By using the terms "infectious period" and "infection rate" we already gave away the relation with epidemic spread. We need this extension because in real-life response to an infection may differ between individuals. One of the things that may be different is the period that an individual stays infectious. In the epidemic model corresponding with locally dependent percolation, these infectious periods are assumed to be i.i.d.

Now define $p := \mathbb{E}(1 - e^{-\tau\mathcal{I}})$ as the marginal probability that a given bond is open. Kuulasmaa [51] proved that if we compare locally dependent percolation with the same marginal probabilities that edges are open, a fixed infectious period gives a worst case scenario, in the sense that the probability of the origin being in an infinite cluster is maximal. This model corresponds with bond-percolation. In [25] it is shown that the process where all edges going out of a vertex are open with probability p and all of them are closed with probability $1-p$ gives a "best-case" scenario, in the sense that the probability of the origin being in an infinite cluster is minimal.

1.6 An overview of this thesis

The goal of the research presented in this thesis is the development of new stochastic epidemic models and methods, to extend the existing methods and to apply models and methods from other areas of probability theory or more general mathematics to epidemiology. When I started the project The Netherlands had just been hit by several outbreaks of infectious diseases of animals, that have had great impact on animal health and welfare and the economy of the country. During the outbreaks the government and the farmers took measures to stop the epidemic and to save animals. One of the effects of the measures is that the dynamics of the spread changes during the outbreak itself. Of course this is exactly what the measures are for: reducing the probability that the infection is transmitted from an infected farm to a susceptible one. However, this change in the dynamics is hard to incorporate in the existing models in epidemiology. In Chapter 2 a model is developed that can be used to study the real-time dynamics of the spread of an infection in varying environments. This model has been used to analyse the effects of various control measures.

The model proposed in Chapter 2 is mainly on branching processes in varying environments. In that branching model properties of the infection and the parameters that describe the spread of the infection in a non-varying environment are used and it is assumed that these parameters are known. In Chapter 3 the leading question is whether we can obtain the needed parameters from a branching process if we only have restricted observations?

It is proven in [34] that if we only observe the generation sizes, then we can only estimate two parameters of the branching process consistently. This means that even if we observe the process for an infinite period of time, we can only be sure about two of the parameters. In our case we do not observe the generation sizes, but only the individuals that stop being in the process (the farms at which an infection is detected). Surprisingly, with these observations it is possible to estimate three parameters consistently. One drawback of the approach is that two of the three estimators converge very slowly and for practical purposes much too slow. One of the conclusions drawn in this chapter is that more information is needed than only the "detection times" of the farm in order to estimate the parameters in the branching model of infection spread.

In Chapters 2 and 3 real-time processes are studied in large randomly mixing populations. However, for some purposes the assumption of randomly mixing populations is strong and seems to be unrealistic. In Chapters 4, 5 and 6 methods are provided to study the spread of infections on networks. However, the cost of this extra structure is that in the Chapters 5 and 6 we loose the real-time perspective on the epidemic. In the three chapters on infection spread on networks some of the properties of the (social) network on which the infection spreads are captured in the models.

In Chapter 4 the method of pair-approximations is discussed. This method is used in deterministic models and it can roughly be seen as a model of randomly mixing pairs of individuals in stead of randomly mixing single individuals. Although the intuition behind this model seems to be clear, it is hard to make all the assumptions and approximations, usually used in pair approximation models, explicit. In this thesis I look at the approximation from the viewpoint of a probabilist. The model assumptions are made explicit and the approximations used in the existing models are analysed. Furthermore, the dynamics of the expectation of the number of infected individuals are described instead of assuming that this expectation is the same as the actual number of individuals (which is the assumption in most of the deterministic models in epidemiology). Furthermore a new proposal for a useful reproduction number is made.

In Chapter 4 the time dynamics of the epidemic are studied. However keeping track of this time dynamics makes it necessary to make some extra assumptions. In Chapter 5 the real-time perspective is replaced by a generation perspective, where we do not consider the times at which infections took place, but instead of that we use the "infection tree", where connections are the actual infections that took place. Note that for quantities like the probability of a large outbreak and the expected number of individuals that will ultimately be infected, the perspective does not matter. Still, mathematical analysis of the spread of an infection on a general network is far too difficult, therefore we replace the general network by a random network which has some features in common with the original network. We determine a reproduction number and the probability of a major outbreak on the random network.

Replacing networks by random networks for modelling purposes is not new. However, in the existing models on epidemics on networks, the random network that is used has no or only few triangles in it. In real-life triangles

naturally arise in social networks (the friends of my friends are often also my friends). In Chapter 5 random networks are constructed that do have an a priori given expected number of triangles.

Finally, in Chapter 6 the relation between a certain type of random networks (the generalised random graphs) and epidemic spread is established and results known from percolation theory (especially the results of Kuulasmaa [51], see Section 1.5) are generalised to these random graphs in order to give worst case scenarios for epidemics in heterogeneous populations. The generalised random graphs are well fit to model the spread of an infection in a heterogeneous population, where individuals have random "infectivity" and "susceptibility' (the weights of the individuals). This chapter can be seen as an "ansatz" for further research.

1.7 Possible future work in epidemiology

As mentioned above, Chapter 6 gives a foundation for further work on modelling the spread of infections in heterogeneous populations. Especially the important question on the population effect of vaccines which do not give total protection and which cause a change in infectivity, can be tackled. Other future work can be expected from the field of percolation theory.

1.7.1 Long-range percolation

The model described in Section 1.5 is elegant, but it does not give a very realistic description of the spread of an infection. A possible extension of the model is *long-range percolation*. This model has not directly been used in this thesis, but it is useful to compare this section with the generalised random graphs of Chapter 6. While edges in generalised random graphs are open based on the weights of the vertices, in long-range percolation they are open based on the distance between two vertices. We change the terminology a bit by replacing "open" by "present" and "closed" by "absent". Using this terminology makes it clearer that we are actually constructing a random graph.

The construction is done as follows: We start with a vertex set V in some metric space. For simplicity we assume $V = \mathbb{Z}^2$ or more general $V = \mathbb{Z}^d$. For each pair of vertices (v, w) there is an edge joining them with probability $p(\|v-w\|)$, where $\|v-w\|$ denotes the Euclidian distance (any other metric will

do as well). The presence or absence of an edge is independent of the presence or absence of other edges, so we still use product measure. It is natural to assume that $p(x)$ is a decreasing function.

Long-range percolation provides a much richer way of modelling the spread of infections than ordinary percolation. Of course analysis becomes much harder as well. In general it is very hard to answer questions on extinction probabilities. However, questions that are little studied in ordinary percolation arise in long-range percolation. One of those questions is concerned the diameter of large subgraphs of $\mathbb{Z}^d$. Consider the block $B_N := [0, N]^d \cap \mathbb{Z}^d$. Edges are added according to the rules of long-range percolation. The random graph obtained in this way is called G. Define for $v, w \in B_N$, $D(v, w)$ as the *graph distance* or *chemical distance* between v and w, i.e. the minimal number of edges in G that has to be crossed to go from v to w. If there is no path from v to w the chemical distance is said to be infinite. The diameter D_N is defined as

$$D_N := \max_{v,w \in B_N} D(v, w). \tag{1.3}$$

In general one restricts attention to a connected component of the graph, because $D_N = \infty$ is not informative.

Results on D_N and $D(v, w)$ for different regimes of $p(x)$ are given in [15, 16, 18, 24]. Marek Biskup proved the following theorem:

Theorem 1.7.1 *Consider long-range percolation on $\mathbb{Z}^d$. Let $s \in (d, 2d)$. Let $p(x) = 1 - exp[-q(x)]$, with*

$$\lim_{|x| \to \infty} \frac{\log q(x)}{\log |x|} = -s \tag{1.4}$$

and assume that, the random graph G a.s. contains a unique infinite component $\mathcal{C}_\infty$. Then for all $\epsilon > 0$,

$$\lim_{|x| \to \infty} \mathbb{P}\Big(\Delta - \epsilon \leq \frac{\log D(0, x)}{\log \log |x|} \leq \Delta + \epsilon | 0, x \in \mathcal{C}_\infty\Big) = 1, \tag{1.5}$$

where

$$\Delta = \Delta(s, d) = \frac{\log 2}{\log(2d/s)}. \tag{1.6}$$

In an unpublished paper Biskup [19] has proven that (with the same notation as before) if $p(1) = 1$ for all $\epsilon > 0$,

$$\lim_{|N|\to\infty} \mathbb{P}\Big((\log N)^{\Delta-\epsilon} \leq D_N \leq (\log N)^{\Delta+\epsilon}\Big) = 1. \tag{1.7}$$

Here the assumption $p(1) = 1$ makes sure that the graph is connected.

For epidemiological purposes one might be interested in the number of vertices within chemical distance n of the origin. We denote this number of vertices by Z_n. A natural question for epidemiologists is, whether it is possible to have exponential growth of Z_n in n for long-range percolation? Only if that is the case it is sensible to use the concept of R_0 in this setting, because although R_0 is defined as the expected number of individuals an initial individual infects, it is usually interpreted as the base of the exponential growth of the generations in the infection tree, e.g. if the generation sizes grow polynomially, the base of the exponential growth will go to 1 and the only sensible reproduction ratio will be hardly informative. It can be proven that for certain slowly decreasing $p(x)$ with the property that

$$\sum_{v\in\mathbb{Z}^d} p(|v|) < \infty,$$

like $p(x) = x^{-d}(\log x)^{-2}$, there exists $1 < m_1 < m_2$ such that $(m_1)^{-n}\mathbb{E}(Z_n) \to \infty$ and $(m_2)^{-n}\mathbb{E}(Z_n) \to 0$ [71]. However, this does not imply that there exists an $\epsilon > 0$ and an $m_3 > 1$ such that there is a positive probability that $m_3^{-n} Z_n > \epsilon$ for all n. In other words, it is possible to generate a long range percolation cluster such that $\mathbb{E}(Z_n)$ grows exponentially, but it is not yet proven whether it is possible that Z_n grows exponentially for some function $p(x)$ with positive probability. It is also not proven whether or not there is a positive probability that Z_n grows exponentially in the regime that is studied in the papers of Biskup [18, 19], where $d < s < 2d$. Simple arguments from spatial branching processes show that for $s > 2d$, the vertex at graph distance n with the largest Euclidean distance to the origin grows with probability 1 slower than a^n for every $a > 1$.

The results of the previous paragraph are interesting in interpreting the results of Chapters 4 and 5, because we made plausible that in long-range percolation models the reproduction ratio is determined by the tail of $p(x)$, while the number of triangles in the long-range percolation graph is mainly due

to short edges and therefore mainly depends on $p(x)$ for small x. However, we can justify the use of the models in Chapters 4 and 5 by the fact that although there is a spatial component in social networks, they are not purely spatial.

1.7.2 Continuum percolation

In this thesis the idea of percolation is introduced by the example of the spread of a disease in an orchard, where the locations of the trees form a two-dimensional grid. However, we may ask what can be done if an infection spreading in a forest is considered? In a forest the positions of the trees are not in any sense "grid like", but these positions seems to be random.

A mathematical model to study the spread of an infection in a forest is continuum percolation [55]. Like in locally dependent percolation, there are two levels of randomness. First, the position of the trees are assumed to be generated by a random point process in $\mathbb{R}^d$ (e.g. a Poisson process in $\mathbb{R}^2$), after that connections are made between pairs of trees based on the distance between the trees. There are two major models in continuum percolation, the Boolean model and the random connection model. In the Boolean model i.i.d. radii are assigned to the points of the underlying point process and there is a connection between a pair of points v and w if their distance is less than $r_v + r_w$, where r_v and r_w are the radii of the points v and w.

In the random connection model the vertices v and w are connected with probability $p(\|v - w\|)$ where $p(\cdot)$ is a decreasing function. The connections are made independently of each other.

Continuum percolation seems to be a useful tool to get insight in infection spread. However, only very few results are known for these models and the results that are obtained can hardly be used in a quantitative setting. One of the open problems that is interesting for epidemiological purposes is the question: "How do percolation probabilities depend on the point process that is used for the continuum percolation?" More specific: instead of using a Poisson point process, one might want to use other Markov point processes [52] to model the positions of the individuals that can spread an infection. Relevant questions are: "Is it possible to give a worst-case point-process like we can give a worst case infectious period for locally dependent percolation?" and "Do attracting or repulsive point processes make percolation more or less probable?"

1.8 List of publications

(a) Trapman, P.; Meester, R. and Heesterbeek, J.A.P. (2004), A branching model for the spread of infectious animal diseases in varying environments, *Journal of Mathematical Biology* **49** 553-576.

(b) Meester, R. and Trapman, J.P. (2005), Estimation in Branching Processes with restricted observations, *submitted:* available at `http://www.math.vu.nl/~ptrapman/ebartikel221205.pdf`

(c) Trapman, J.P.(2006), On analytical approaches to epidemics on networks, *submitted:* available at `http://www.math.vu.nl/~ptrapman/TPBpaper.pdf`

(d) Trapman, J.P.(2006), A reproduction number for epidemics on networks, *preprint.*

Chapter 2 is based on (a), Chapter 3 is almost identical to (b), Chapter 5 is based on (c) and Chapter 4 is based on (d).

Chapter 2

A branching model for the spread of infectious animal diseases in varying environments

2.1 Introduction

Recent outbreaks of infectious diseases of animals (e.g. classical swine fever (CSF), foot and mouth disease (FMD) and Avian Influenza (AI)) in Western Europe have had great impact on the economy, public life and animal health and welfare in the countries involved. During such an outbreak one would like to be able to compare the effectiveness of proposed control measures in, for example, their ability to reduce the expected final size and the expected duration. Typical for the strategies aimed at stopping outbreaks of important diseases of farm animals, is that infected herds are removed from the population by culling upon detection. A second characteristic is that due to increasing quantity and quality of the imposed control measures, the environment that the infectious agent experiences, is changing. By this we mean that consecutive measures can make, for example, contact opportunities between herds different in different phases of the outbreak, or can make the infectious period, or rate with which infectivity is produced, differ for farms infected at different times.

Mathematical methods for computing outbreak characteristics such as expected final size and expected duration in most cases assume a constant environment in that the control measures are not compounded in time and do not lead to changes in the rates that govern epidemic spread (see e.g. [27]). In this chapter we aim to develop stochastic methods, based on the theory of branching processes, which allow us to compare the effectiveness of control strategies during such outbreaks where the environment is varying because of changes in subsequent measures of control.

Much work has already been done to describe the spread of classical swine fever (see e.g. [49, 56, 69] and Chapter 3) and foot and mouth disease (see e.g. [29, 30, 40, 42]). In this chapter, we will model the spread of infections in a much more analytic way than is done in earlier models [29, 30, 42, 49]. We use an iterative method that computes properties of the spread, like the probability of a major outbreak of the infection and the final size of an epidemic (i.e. the total number of infected herds), very efficiently. Furthermore, we can derive some properties of the duration of the epidemic. We allow for different types of herd.

In our model, it is essential that once the infection in a herd is detected, the whole herd will be culled. Therefore, our main interest is the number of infected herds. However, the number of infected animals in an infective herd is important for determining the infectivity of a herd and the distribution of the detection times. We model the spread of the infection at two levels, namely the spread of the infection within a herd and the spread of infection between herds; both are described by a stochastic process. The distribution of the number of infective animals within a herd is incorporated in the model for the spread of the infection among herds.

We use a special branching process to describe the spread of the infection among herds. The parameters of this branching process depend on the time since the infection of the herd and on the environment, which is determined by the real time (as opposite to the time since the infection of a specific herd). Using branching processes to describe epidemics is of course not new (see e.g. Section 1.4 and [39]), but no theory exists that gives short-term predictions (as opposed to asymptotics) for general branching processes in varying environments, with an age-dependent "birth rate".

For computations it is necessary that after a certain moment the environment is constant. To achieve this we assume that after some time no new measures will be taken and the effects of all measures taken in the past will either be constant or absent. In other words, although the values of the parameters may differ from those before certain measures were taken, the values are assumed to be constant after a given moment in time.

Our model can be used to predict the effects of various control measures and strategies during an ongoing outbreak. Meester et al. gave a method to estimate the parameters from the data available during an outbreak ([56] and Chapter 3). We consider measures like single vaccination of all herds of a certain type, a total transport ban or killing of animals just after birth. Some of these measures cause a varying environment. The fraction susceptible animals in a herd or the fraction susceptible herds of the total number of herds may be varying and so the infection rate may vary in time too. In fighting outbreaks, additional measures will be implemented as soon as present measures turn out to be insufficient. A change in measure will change the values of the parameters. We assume that the changes in the environment are deterministic.

We develop the theory using a classical swine fever outbreak in herds of pigs as motivating example throughout. In this chapter we use data of classical swine fever (CSF) for our computations. These are the same as the input data Klinkenberg et al. used in [49].

2.2 Spread of the infection within a herd

2.2.1 The model

As mentioned in the introduction, we first need to model the spread of the infection within one herd, since the infectivity of an infective herd and also the time at which the infection is detected in a certain herd, depend on the number of infected animals in that herd. We use t for the time elapsed since the first measures were implemented. The variable τ is used for describing the spread of the infection within the herds, it is the time since infection of a particular herd and therefore relative. The τ-clock starts ticking at the moment the first animal in the herd becomes infected. From now on, we will call τ the "infection-age" or "age" of the herd.

We use only four types of disease-related parameters:

λ**:** The infection rate of individual animals within a herd.

μ**:** The recovery rate of individual animals in a herd.

α**:** The per capita detection rate of infected animals within a herd.

β**:** The rate at which one infected animal infects susceptible herds.

In this section we only use the first two, in the next section we also use α and β.

We assume that as soon as an infection at a particular herd is detected, the whole herd will be culled instantaneously. Further we assume that the rate of detection is proportional to the number of infected animals at infection-age τ, $I(\tau)$, i.e. the detection rate is $\alpha I(\tau)$, where α does not depend on the infection-age. We assume that the infection in a herd develops as an autonomous process until detection. The number of infected animals in a herd therefore depends only on the infection-age τ, and not on the absolute time t.

We describe the number of infective animals by an ordinary birth and death process [32]. Writing $p_i(\tau) = \mathbb{P}(I(\tau) = i)$ for the probability of i infective animals in a herd at infection-age τ and δ_{ij} for the Kronecker Delta function (i.e. $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise), we have (see e.g. [32]):

$$\begin{aligned}
p_i(0) &= \delta_{i1}, \\
\frac{dp_0(\tau)}{d\tau} &= \mu p_1(\tau), \\
\frac{dp_1(\tau)}{d\tau} &= -(\lambda+\mu)p_1(\tau) + 2\mu p_2(\tau), \\
\frac{dp_i(\tau)}{d\tau} &= -i(\lambda+\mu)p_i(\tau) + (i-1)\lambda p_{i-1}(\tau) + (i+1)\mu p_{i+1}(\tau) \qquad \forall\ i \geq 2.
\end{aligned}$$

Solving these differential equations leads to:

$$\begin{aligned}
p_0(\tau) &= \frac{e^{r\tau}-1}{Re^{r\tau}-1}, \\
p_i(\tau) &= (1-p_0(\tau))(1-Rp_0(\tau))(Rp_0(\tau))^{i-1} \qquad \forall\ i \geq 1,
\end{aligned}$$

where $r = \lambda - \mu$ and $R = \frac{\lambda}{\mu}$ is the reproduction ratio. We assume $\lambda > \mu$, otherwise the infection will only cause a minor outbreak within a herd and between herd infections are very rare.

Hence, conditioned on the event that the epidemic in a herd does not go extinct before age τ, i.e. $I(\tau) > 0$, $I(\tau)$ is geometrically distributed with parameter $(1 - Rp_0(\tau)) = \frac{R-1}{Re^{r\tau}-1}$, which is small for large τ.

The ceiling of an exponentially distributed random variable with parameter x is a geometric random variable with parameter $1 - e^{-x}$. For small x we can use the approximation $1 - e^{-x} \approx x$. Furthermore, on the set $\{I(\tau) > 0\}$, $I(\tau)$ is large with high probability for large τ. Therefore, for large τ, $I(\tau)$ can be approximated by $\bar{H}(\tau)$, where $\bar{H}(\tau)$ is an exponential random variable with parameter $\frac{R-1}{Re^{r\tau}-1}$.

Further, if X is exponentially distributed with parameter x, than cX is exponentially distributed with parameter x/c. Therefore, $\bar{H}(\tau)$ is distributed as $H(e^{r\tau} - \frac{1}{R})$, where H is an exponential random variable with parameter $\frac{R-1}{R}$. For large τ, the term $\frac{1}{R}$ is negligible compared to $e^{r\tau}$. So we use the approximation:

$$I(\tau) \approx He^{r\tau}, \tag{2.1}$$

where the approximation is in the distribution sense.

We can interpret this approximation as follows, using the terminology of branching processes (Section 1.4). The random variable H is due to the random character of the start of the outbreak in a herd, when only a few animals are infective. If the disease will not go extinct, after the initial phase many animals are infected in the herd, each of which causes an expected offspring of e^r new infections per time unit. Hence the growth rate is eventually almost deterministic due to the law of large numbers, but the initial dynamics of the spread are viewed as random. We also use this approximation for small τ.

We use the word infective for animals or herds that are able to spread the infection. We use the notation $I(\tau; h)$ for the number of infective animals in a particular herd, with $H = h$ given.

2.2.2 Discussion of the within-herd model

1. We assume that the infection and recovery rate of individual infected animals are independent of time and age.
2. It is not necessary to use a birth and death process to describe the spread within a herd. We may also use other processes, for example the one dimensional nearest neighbour contact process (in this chapter called the contact process). In this model we think of a situation where all animals are posi-

tioned in a row and do not change position. Each animal can only infect its two nearest neighbours. We assume that recovered animals with two infective neighbours are re-infected immediately. Therefore, we only consider the animals at the edge of a row of infective animals. We still assume that the recovery rate is μ. An infective animal infects each of its susceptible neighbours with rate λ. Now we have:

$$\begin{aligned} p_i(0) &= \delta_{i1}, \\ \frac{dp_0(\tau)}{d\tau} &= \mu p_1(\tau), \\ \frac{dp_1(\tau)}{d\tau} &= -(2\lambda+\mu)p_1(\tau) + 2\mu p_2(\tau), \\ \frac{dp_i(\tau)}{d\tau} &= 2(-(\lambda+\mu)p_i(\tau) + \lambda p_{i-1}(\tau) + \mu p_{i+1}(\tau)) \qquad \forall\ i \geq 2. \end{aligned}$$

We denote the number of infectives by $I(\tau)$. One can show that $\mathbb{E}(I(\tau)|I(\tau) > 0) = 2r\tau + o(\tau)$. Further we can show that $\mathrm{Var}(I(\tau)|I(\tau) > 0) = o(\tau^2)$. This implies that for $I(\tau) > 0$ we have $\frac{I(\tau)}{\mathbb{E}(I(\tau)|I(\tau)>0)} \to 1$. So, in contrast to the birth and death model, we may approximate the random variable $I(\tau)$ for large τ by a deterministic variable: $2r\tau$.

3. For infections in different types of herds, we may need different models for the within-herd spread. Some animals live in small compartments in a row. Other animals live in large herds where all susceptible animals are equally likely to be infected by one infective animal. For the former we use the contact model and for the latter we use the birth and death model.

4. In the original model, the number of animals in one herd is assumed to be very large compared to the number of infected animals. Therefore, we assume that the contact rate between infected and susceptible animals is constant, and hence so is the birth rate. From data of outbreaks of CSF in the past, we can see that the number of infected animals until detection of the infection within the herd, is small compared to the total number of animals in the herd, so these assumptions seem justified in this case [69].

5. The within-herd infection and recovery parameters λ and μ can be measured experimentally or from data of past or on-going outbreaks.

   The parameter α depends on the development of symptoms of infected animals and how attentive the farmers are. Therefore, in reality this α will

change at the first detection of the infection in the country (or in neighbouring countries), due to higher awareness of farmers and veterinarians. It is very difficult to estimate α for the period before the time of the first detection. For the time after the first detection, we can estimate α from data of past outbreaks in the same area or try to estimate this parameter during the ongoing epidemic. Using data from past outbreaks is dangerous, because the characteristics of the virus and of the farming practice may have changed. Estimation of the parameters, during an on-going outbreak is done by Meester et al. [56] and in Chapter 3. This method has some problems, e.g. the time necessary to get enough data for a reliable estimate. Another problem is that in [56] and in Chapter 3 the infectivity of a herd does not depend on the "age" of the herd. This independence of age is essential for the estimations made.

For our model it is not necessary that α is constant in time. It is possible to extend the model and use $\alpha(t)$ instead of α. For varying values of α we need to use the model for varying environments.

6. We assume that culling is the only measure which influences the within-herd spread of the infection. Vaccination is assumed to show no effect on the spread in the herd. For CSF this assumption is justified by the fact that vaccination will lead to immunity only after two weeks. Therefore, during this first two weeks the spread of the infection is not affected by this measure. For the time after these two weeks, we assume that we can use the same speed of propagation of the infection, for computational reasons. The approximation leads to results that are conservative, that is, too pessimistic.

7. We do not take characteristics of individual animals into account that might cause the individuals to differ in infectivity, susceptibility or contact pattern. Often age and type of species can have a substantial influence. For CSF this implies we do not distinguish between the age of animals in one herd. The detection rate and infectivity of herds with many young animals does not significantly differ from the detection rate and infectivity of herds with especially older animals [49, 69].

8. We also use the approximation $I(\tau) \approx He^{r\tau}$ for small τ on $\{I(\tau) > 0\}$. This is not correct, but due to the small number of infective animals at small τ, the probability that the disease is detected at small τ, will be small too. We will later see that the number of infections in that period is small too, so

the overall influence of the between-herd-events (infections and detections) while the herd is "young", is small.

9. For classical swine fever, the spread of the infection seems to be well described with a birth and death model. For this reason the parameter values for the contact model were not estimated and we did not do computations in this model. We included the description of the contact model because other infectious diseases may spread according to this model. We note that including this model makes computations easier than in the case of the birth and death model, because we do not have to deal with a random variable for the within-herd spread. So if we have estimated the values of the parameters, we can easily estimate the same properties as we do for the birth and death model.

2.3 Spread of the infection between herds

2.3.1 The model for non-varying environments

In this section, we consider classical swine fever as a concrete example.

We distinguish between two types of farms (or herds): *multipliers* (m) are roughly speaking farms where young piglets are born, and *finishers* (f) are farms that buy piglets and fatten them. The parameter p_m denotes the fraction multipliers of the total number of herds and p_f is the fraction finishers. We assume that within either of these types, a birth and death model (with the same parameters for both types) describes the within-herd spread. The infectivity per non-transport contact of both types of herds develops in the same way too. However, transport contacts are only allowed from multipliers to finishers. Therefore, we take a larger infection rate for contacts from multipliers to finishers.

We define $A_\xi(t, \tau; h)$ as the infectivity of a herd (the rate at which contacts are made with other herds) at time t, while the herd was infected τ time units ago and with $H = h$, where ξ is a two-dimensional vector, denoting the two types of herds involved in the contact, so that ξ can be ff, fm, mm or mf. When no measures are implemented, $A_\xi(t, \tau; h)$ is proportional to the number of infected animals in the herd:

$$A_\xi(t, \tau; h) = \beta_\xi I(\tau; h).$$

Here β_ξ is a constant depending only on ξ. Because the non-transport contacts all happen at the same rate we can define $\beta_m := \beta_{fm} = \beta_{mm}$. Further we write $\beta_f := \beta_{ff}$ and β_{mf} is β_f plus some additional term for infections caused by transport of piglets from multipliers to finishers. Because all non-transport infections happen at the same rate the ratio $\beta_m : \beta_f$ is exactly the ratio of the multipliers to the finishers. Note that the infectivity does not depend on the absolute time t (if there are no measures implemented).

We define β by $\beta := \beta_m + \beta_f$. So $\beta_m = p_m\beta$ and $\beta_f = p_f\beta$. In order to consider transport contacts we also define $\beta_{mf} = p_f\beta + \beta_{tr}$, where β_{tr} is the proportionality factor of the part of the infection rate that is due to transport contacts. There is no term p_f in front of β_{tr}, because all transport contacts are from multipliers to finishers. We assume that the total number of herds is very large; in our computations, we assume it infinite.

We already know from the definition of α in the previous section, that the detection-rate is given by $\alpha I(\tau; h) = \alpha h e^{r\tau}$. From this we can deduce (See page 113 of [27]), for given h, the probability that an infected herd of age τ is not yet detected, $p_{nd}(\tau; h)$:

$$p_{nd}(\tau; h) = e^{-\int_0^\tau \alpha h e^{rs} ds} = e^{-\frac{\alpha}{r} h(e^{r\tau}-1)} \tag{2.2}$$

In the case where no measures are implemented, the expected number of infected multipliers by one infective multiplier up to age τ, for given h, $\mu_{mm}(\tau; h)$, is given by:

$$\mu_{mm}(\tau; h) = \frac{\beta_m}{\alpha}(1 - e^{-\frac{\alpha h}{r}(e^{r\tau}-1)}). \tag{2.3}$$

To see this, we first consider only one type of herd. We prove the following proposition:

Proposition 2.3.1 *Suppose that the infection and detection rate are proportional to the number of infective animals in a herd and that the environment is non-varying. Then the distribution of the size of the progeny of an infective herd does not depend on the number of infective animals in that herd at the start of the process.*

Proof. The infection and the detection rate are proportional to the number of infective animals in a herd. The ratio of infection rate and detection rate is given by $\beta : \alpha$. We call the detection of the herd and the infections by this herd *events*. The probability that an event is an infection is $\frac{\beta}{\alpha+\beta}$ and a detection $\frac{\alpha}{\alpha+\beta}$. Therefore, the number of events, including detection, is described by a geometric random variable with parameter $\frac{\alpha}{\alpha+\beta}$. Therefore, the direct offspring distribution does not depend on the size at the start of the process. Moreover, the same holds for the offspring of this direct offspring. □

In the same way we can prove that the size of the future offspring of an infective herd is independent of its age.

To prove (2.3), consider two types of herds. Let N_{mm}, be the number of multipliers infected by one infective multiplier. We note that $N_{mm} + 1$ is distributed as a geometric random variable with parameter $\frac{\alpha}{\alpha+\beta_m}$ (1 is added because the final event will be the only detection, and the number of events is described by a geometric random variable, where an event is defined as in the proof of the proposition). So the expected number of future infections of an infective multiplier is $\frac{\beta_m}{\alpha}$, at all times. The probability that an infective herd is not yet detected at age τ is given by $e^{-\frac{\alpha}{r}h(e^{r\tau}-1)}$. From this and Proposition 2.3.1 we deduce that the expected number of infections after age τ is given by $\frac{\beta_m}{\alpha}e^{-\frac{\alpha}{r}h(e^{r\tau}-1)}$. By subtracting the expected number of infections after age τ from the expected total number of infections by one herd we get the expected number of infections until age τ. In the same way we can deduce

$$\begin{aligned}
\mu_{mf} &= \frac{\beta_{mf}}{\alpha}(1 - e^{-\frac{\alpha h}{r}(e^{r\tau}-1)}),\\
\mu_{fm} &= \frac{\beta_m}{\alpha}(1 - e^{-\frac{\alpha h}{r}(e^{r\tau}-1)}),\\
\mu_{ff} &= \frac{\beta_f}{\alpha}(1 - e^{-\frac{\alpha h}{r}(e^{r\tau}-1)}).
\end{aligned}$$

Now consider the probability p^f_{kl} that a finisher infects k multipliers and l finishers. All events (infections of multipliers, infections of finishers and detection of an infected herd) happen at a rate proportional to the number of infective animals in the infective herd. Therefore, the proportions of the rates stay the same. First, we only consider infections and detections, and we do not yet consider the different types of herds infected. Detection occurs with rate $\alpha h e^{r\tau}$ and infection occurs with rate $\beta_m h e^{r\tau} + \beta_f h e^{r\tau} = \beta h e^{r\tau}$. As in

the proof of Proposition 2.3.1 we can describe the total number of events, $\bar{D}$ say, by an ordinary geometric random variable with parameter $\frac{\alpha}{\alpha+\beta}$. So the probability that $n+1$ events occur, i.e. n herds are infected by one finisher, is

$$\mathbb{P}(\bar{D} = n+1) = \Big(\frac{\alpha}{\alpha+\beta}\Big)\Big(\frac{\beta}{\alpha+\beta}\Big)^n$$

If in total n herds are infected by one finisher, we know by the lack-of-memory property of the infection process that the number of infected multipliers, N_{fm}, is binomially distributed with parameters n and $\frac{\beta_m}{\beta}$. That is,

$$\mathbb{P}(N_{fm} = k|\bar{D} = n+1) = \binom{n}{k}\Big(\frac{\beta_m}{\beta}\Big)^k\Big(\frac{\beta_f}{\beta}\Big)^{n-k}.$$

Note that

$$\begin{aligned} p^f_{kl} &= \mathbb{P}(N_{fm} = k, N_{ff} = l) \\ &= \mathbb{P}(N_{fm} = k|\bar{D} = k+l+1)\mathbb{P}(\bar{D} = k+l+1) \\ &= \binom{k+l}{k}(\frac{\beta_m}{\beta})^k(\frac{\beta_f}{\beta})^l(\frac{\alpha}{\alpha+\beta})(\frac{\beta}{\alpha+\beta})^{k+l}. \end{aligned}$$

The generating function of $\{p^f_{kl}\}$ is now given by

$$g^f(s_1, s_2) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} p^f_{kl}s_1^k s_2^l = \frac{\alpha}{\alpha+\beta-\beta_m s_1-\beta_f s_2}.$$

We can deduce

$$g^m(s_1, s_2) = \frac{\alpha}{\alpha+\beta+\beta_{tr}-\beta_m s_1-(\beta_f+\beta_{tr})s_2}$$

in the same way. Note that we did not need the distribution of the random variable H, we only needed the proportionality factors of the parameters.

In the special case where infection and detection rates are proportional and these rates are known for the time before the first detection, we can also give the distribution function of the number of infective herds at the time of the first detection. This distribution function is the same as the distribution function of the number of direct infections by one herd, because it still holds that an event is a detection with probability $\frac{\alpha}{\alpha+\beta}$ and the probabilities of

infections of finishers and multipliers are respectively $\frac{\beta_f}{\alpha+\beta}$ and $\frac{\beta_m}{\alpha+\beta}$.

We are also interested in the probability that a herd infects k multipliers and l finishers, before the herd reaches age τ. We consider a finisher. We write $\mathbb{P}(N_{fm}(\tau;h)=k, N_{ff}(\tau;h)=l)$ for this probability when the detection age τ_d and $H=h$ are given. The infections before age τ_d occur independently of each other and have the lack of memory property. Therefore, for $\tau \leq \tau_d$ the times of infections of the different types of herds are described by an inhomogeneous Poisson-process with rate $\beta_\xi h e^{r\tau}$. So the number of infections until age τ is Poisson distributed with parameter

$$h\beta_\xi \int_0^\tau e^{rs} ds = h\beta_\xi (e^{r\tau}-1).$$

Now we see:

$$\mathbb{P}(N_{fm}(\tau;h)=k, N_{ff}(\tau;h)=l) = \frac{(h\frac{\beta_m}{r}(e^{r\tau}-1))^k}{k!} \frac{(h\frac{\beta_f}{r}(e^{r\tau}-1))^l}{l!} e^{-h\frac{\beta}{r}(e^{r\tau}-1)}.$$

Note that now we know the life length distribution, the expected number of infections by a herd up to age τ and the underlying Galton-Watson process of the branching process, we can use the general theory of branching processes to determine some other properties, like the expected duration of the outbreak. (Chapters 2 and 6 of [39]).

2.3.2 Discussion

1. We use a branching idea to describe the spread of the infection among different herds. For this approach, we assume that one herd has contacts with many other herds. In our model we do not take spatial distribution of the herds into account. As long as there are many herds in an area, the local exhaustion of susceptible herds can be ignored. If the outbreak is in an advanced state, however, local depletion of susceptible herds (either by the epidemic progression or by so-called pre-emptive culling) will certainly play a role. As long as there are relatively few infected herds in a neighbourhood, (i.e. a group of herds that have contacts with each other) we assume that the infection rate does not directly depend on this number of infected herds. Measures like ring-culling (culling of all herds within a certain distance of an infected herd) and ring vaccination cannot be considered in our model.

2. We distinguish between two types of contacts, transport contacts and indirect contacts. We have assumed that the indirect contacts are the same for all herds. Transport contacts are only from multipliers to finishers. Indirect contacts include transport of the virus by wind, by visits from an infective herd to a susceptible herd etc. Transport contacts are transports of infected animals from a multiplier to a finisher.

3. We simplify the model by assuming that the measures are implemented at the time of the introduction of the virus in the population (or, in other words, that the infection is detected immediately). It would be desirable to implement the measures from the moment of first detection. It is difficult, however, to estimate parameters like α and β for the time before the first detection, when awareness is not yet heightened by announcement of the outbreak, and when increased hygienic measures on farms are not yet taken. If we know the value of the parameters of the model for the period between introduction and first detection, we are able to estimate the probability of extinction and the expected final size of such a situation. However, these calculations still require that we use the information from our simplified model, where measures are started directly upon introduction.

4. The function $\mu_{mm}(\tau; h)$ is not required for our calculations of the final size and the probability of extinction. The reason why we include this function in our chapter is that this function is interesting in its own right. It gives the expected offspring τ time units after an infective herd is infected itself. In practical applications, it is possible that by contact tracing some herd is suspected of being infected τ time-units ago. This $\mu_{mm}(\tau; h)$ is then the expected number of infected multipliers by the suspected herd, if that suspected herd is a multiplier and if it is really infected.

5. The proportionality of the infection rate to the detection rate is essential in this section. Due to this property, we can find an underlying ordinary Galton-Watson process. Without this proportionality, we may ask whether it is possible to give a nice expression for the generating function? We also lose the independence of h in the generating function, which gives some computational problems.

6. Knowing the generating function of the number of infected herds at the first detection is important, because if in some way it is possible to estimate parameters for the time before the first detection, we can use the

distribution function of the number of infective herds at the time measures are implemented. We do not have the 'ages' of the infective herds at the first detection, but we can find a worst-case-scenario, by finding what age of a herd at $t = 0$ leads to the biggest offspring. If the detection and infection rate does not depend in the same way on the number of infected animals, we cannot deduce the distribution function at the time of the first detection in this way.

2.4 The model for varying environments

2.4.1 The approach

In the previous section, we considered the model for the spread of an infectious disease in non-varying environments. We heavily used the proportionality of the infection and detection rate of herds. This proportionality does not hold in a varying environment. Therefore, we need a different approach.

We consider only one type of herd; the multi-type model is a straightforward generalisation of this. We assume that the spread within a herd is not influenced by the state of the environment. The detection rate only depends on the number of infected animals in a herd and is written as $\alpha I(\tau; h)$, where $I(\tau; h)$ is the number of infected animals as described in Section 2.2. In this section, we assume that the within-herd spread is described by a birth-and-death process. So $I(\tau; h) = he^{r\tau}$. We can easily deal with other descriptions of the within-herd spread.

The infection rate may change due to control measures. Some measures lead to an infection rate that only changes finitely many times, other measures cause a continuously varying infection rate. We assume that in either case the environment is constant after some given time, $t = T$. That is the moment that measures have no more added value and therefore do not lead to new changes in the values of the model parameters. The infection rate is given by

$$A(t, \tau; h) = \beta\phi(t)I(\tau; h),$$

where $\phi(t)$ describes the effects of the measures implemented. We assume that $\phi(t)$ is deterministic. Because the environment does not influence the within-herd spread, $\phi(t)$ does not depend on the age τ. For the rest of this section it is assumed that $\phi(t) = 1$ for $t \geq T$.

The assumption of constant environment after some given time is often realistic. With a vaccination for instance, we know the moment that all vaccinated animals no longer live. In other control measures, like a transport ban, we can vary the time T and compute the effects.

We are looking for the probability of extinction of the infection, the expected final size and a generating function for the number of infected herds at a certain moment. To do this we use a discrete approximation of $\phi(t)$, so we have only a finite number of changes. Because the environment is non-varying after a certain moment, we can use the ordinary theory of branching processes to get all interesting properties for herds infected after $t = T$. By using backward iteration we will find the relevant properties of the epidemic for a herd infected in another interval, because these properties only depend on what happens in the intervals after the interval of infection. We can compute these properties for a herd that was already infected at $t = 0$.

2.4.2 The probability of extinction

Because an infected herd will almost surely be detected in finite time, the probability of extinction of the "progeny" of an infected herd, x, is equal to the probability of extinction of the progeny of all the herds infected by x. The probability of extinction of the "progeny" of a herd infected at time t, is denoted by $q(t)$. So:

$$q(t) = \mathbb{E}(\prod_i q(t_i)) \tag{2.4}$$

(see Section 1.4). Here, the times t_i are the times that herds are infected by the herd infected at time t; the expectation is over these random times of infection. The empty product is defined as 1.

In order to make computations possible we use a discrete time approximation. We divide the positive real line into $N+1$ intervals, labeled $1, 2, \cdots, N+1$. The interval $(0, T]$ is divided into N intervals of equal length, where t is in interval i, if $t \in ((i-1)\frac{T}{N}, i\frac{T}{N}]$, $1 \leq i \leq N$. The final interval $N+1$ is (T, ∞); in this final interval all the model parameters are constant. The time $t = 0$, the moment of the first infection, is not included in one of the intervals, we treat this point in our notation as interval 0.

If the function $\phi(t)$ is discontinuous we may choose another discretisation, so that the discontinuities are on the boundaries of intervals. It is not essential

that all intervals have the same length.

For t in interval i, $1 \leq i \leq N$, $q(t)$ and $\phi(t)$ are approximately constant. In our "discrete time model" we will write $q(i)$ and $\phi(i)$ for the value of q and ϕ in interval i. For instance, we can take $\phi(i)$ to be the value of $\phi(t)$ at the midpoint of interval i. Because we accept discontinuities in $\phi(t)$, this function may differ significantly for neighbouring intervals.

Now, with $n(i,j)$ the number of infections in interval j, due to one herd infected in interval i, we can write (2.4) as:

$$q(i) = \mathbb{E}\Big(q(i)^{n(i,i)} q(i+1)^{n(i,i+1)} \cdots q(N+1)^{n(i,N+1)}\Big), \tag{2.5}$$

Now, the expectation is over the numbers of infections in each interval. $q(0)$ is the probability of extinction of the progeny of the herd infected at time $t = 0$. We assume that all infections and detections take place on the midpoint of the interval, except of course for the events in the final interval, because there we can compute everything explicitly.

Given that a herd is not yet detected and $H = h$, the probability of infection in a certain interval by that herd does not depend on the number of infections in previous intervals. Consider a herd infected in interval i, denote by $\{D(i) = k\}$ the event that this particular herd is detected in interval k. After detection no further infections occur. Define $n(i,l,k;h)$ as the number of infections in interval l due to a particular herd infected in interval i, while $H = h$ and the interval of detection k, are given. Now, by using independence we have (writing P_H for the distribution function of H.)

$$\begin{aligned}
q(i) &= \int \sum_{k=i}^{N+1} \mathbb{E}(q(i)^{n(i,i,k;h)} q(i+1)^{n(i,i+1,k;h)} \cdots \\
&\quad \cdots q(k)^{n(i,k,k;h)} | D(i) = k) \mathbb{P}(D(i) = k | H = h) dP_H \\
&= \int \sum_{k=i}^{N+1} \mathbb{E}\left(\prod_{l=i}^{k} (q(l)^{n(i,l,k;h)} | D(i) = k) \right) \mathbb{P}(D(i) = k | H = h) dP_H \\
&= \int \sum_{k=i}^{N+1} \prod_{l=i}^{k} \mathbb{E}(q(l)^{n(i,l,k;h)} | D(i) = k) \mathbb{P}(D(i) = k | H = h) dP_H.
\end{aligned}$$

Here the expectation inside the integral depends on h, contrary to the non-varying environment case. We condition on the event $\{H = h\}$, which is allowed by Section 4.6 of [32]. Note that the number of infections in a certain

interval is not independent of the interval of detection. It does not make any difference for the number of infections whether a herd is detected shortly after the considered interval or very long after that, but it does make a difference whether the detection is in the considered interval itself. So $p(n;i,l,k;h) := \mathbb{P}(n(i,l,k;h) = n)$ depends on the detection interval k. For computational reasons we suppose in this discrete model that for $i \leq N$, $p(0;i,i,k;h) = 1$. We write $p_{det}(i,k;h)$ for $\mathbb{P}_h(D(i) = k)$. We are interested in the probability of extinction of a herd infected at time 0, $q := q(0)$. We can easily compute all these probabilities in our model.

Because we assume a birth-and-death process for the within-herd spread, we have to take the random character of H into account. Doing this leads to the following formulae:

$$q(i) = \sum_{k=i}^{N+1} \int_0^\infty \frac{R-1}{R} e^{-h\frac{R-1}{R}} (\mathbb{E}(q(i)^{n(i,i,k;h)}) \cdots \mathbb{E}(q(k)^{n(i,k,k;h)})) p_{det}(i,k,h) dh.$$

Here we have conditioned on the time of detection and on h. We can rewrite this formula as:

$$q(i) \quad = \quad \sum_{k=i}^{N+1} \int_0^\infty \frac{R-1}{R} e^{-h\frac{R-1}{R}} (\prod_{l=i+1}^{k} (\sum_{n=0}^{\infty} p(n;i,l,k;h) q(l)^n)) p_{det}(i,k;h) dh.$$

Here $q(i)$ depends on $q(l)$ for $l > i$. As mentioned before, we can compute $q(N+1)$ by the ordinary theory of branching processes. We use backward iteration to compute $q(0)$.

Note that we also know the probability of extinction for herds infected after time $t = 0$. If we estimate (for example by contact tracing) the moment of infection of a certain herd, we can use this information to improve predictions.

2.4.3 The expected final size of the epidemic

In almost the same way as we calculated the probability of extinction, we can calculate the expected final size of the epidemic, i.e. the expected total number of infected herds, G. This will also be done for only one type of herd.

If $q < 1$, there is a positive probability for the final size to be infinite and therefore the expected final size will always be infinite; so to have the possibility of a finite expected final size we assume $q = 1$. We denote by $G(t)$

the size of the progeny of a particular herd, infected at time t (including this ancestor) and write $G(0) = G$. The expected number of herds in the progeny of a given infected herd, including that infected herd itself, is 1 plus the sum of the expected size of the progenies of all herds directly infected by this herd, i.e.

$$\mathbb{E}(G(t)) = 1 + \sum_i \mathbb{E}(G(t_i)),$$

where the t_i's are again the random times at which infections by the element, infected on time t, occur. The empty sum is defined to be equal to 0.

In the same way as we deduced the formulae for $q(i)$ we deduce the formulae for $\mathbb{E}(G(i))$ in the discrete approximation. We write $\bar{G}(i)$ for $\mathbb{E}(G(i))$:

$$\begin{aligned}\bar{G}(i) &= 1 + \sum_{k=i}^{N+1} \int_0^\infty \frac{R-1}{R} e^{-h\frac{R-1}{R}} \mathbb{E}(n(i,i,k;h))\bar{G}(i) + \cdots \\ &\quad \cdots + \mathbb{E}(n(i,k,k;h))\bar{G}(k) p_{det}(i,k;h)dh \\ &= 1 + \int_0^\infty \frac{R-1}{R} e^{-h\frac{R-1}{R}} \sum_{k=i}^{N+1} (\sum_{l=i}^{k} (\sum_{n=0}^{\infty} np(n;i,l,k;h)\bar{G}(l))) p_{det}(i,k;h)dh\end{aligned}$$

For $\alpha > \beta$, $\bar{G}(N+1)$ is known from the ordinary theory of branching processes (see e.g. [39]). For $\phi(N+1) = 1$, $\bar{G}(N+1) = \frac{\alpha}{\alpha-\beta}$.

2.4.4 A generating function for the number of infective herds

We will give the generating function for the number of infective and infected (infective plus removed) herds at any given moment. Here we will consider the generating function for the infective and infected herds at time $t = T$, the time after which the parameter values are considered to be constant.

Remember that the age of an infective herd only influences the number of infective animals in that herd. From Proposition 1 in Section 3.1, we know that the expected direct offspring of a herd, born after $t = T$, is independent of the age of that herd, given the herd is not yet detected at that time. Furthermore we know that the size of the offspring (infected after time $t = T$) of a herd, infective at that time, does not depend on the offspring of other herds that are infective at time $t = T$. So for the distribution of the number of infections after time $t = T$ only the distribution of the number of infective herds at that time is important.

By using Proposition 1 and the theory for ordinary branching processes, we can compute everything we want to know. We will determine the distribution of X, the number of infective herds at time $t = T$. X_i is the number of infective herds at time T, from the progeny of one particular herd infected in interval i. (Here the herd itself is a part of its own progeny.) We denote X_0 by X. We also use Y_i for the number of infected herds in the progeny of a particular herd infected in interval i, that is detected before time $t = T$. The generating function of the distribution of X_i and Y_i is

$$\tilde{g}_i(s_1, s_2) = \mathbb{E}(s_1^{X_i} s_2^{Y_i}) = \int_0^\infty \frac{R-1}{R} e^{-h\frac{R-1}{R}} (\mathbb{E}(s_1^{X_i} s_2^{Y_i}|h))dh.$$

We write $\tilde{g}_i(s_1, s_2; h) = \mathbb{E}(s_1^{X_i} s_2^{Y_i}|h)$.

We again assume that a herd does not infect other herds in the interval wherein it becomes infected itself. So:

$$\begin{aligned}
\tilde{g}_i(s_1, s_2; h) &= \mathbb{E}(s_1^{X_i} s_2^{Y_i}|h) \\
&= \sum_{k=i}^{N+1} p_{det}(i, k; h)\mathbb{E}(s_1^{X_i} s_2^{Y_i}|D(i) = k, h) \\
&= s_2 \sum_{k=i}^{N} p_{det}(i, k; h)\mathbb{E}(\prod_{l=i}^{k} \tilde{g}_l(s_1, s_2)^{n(i,l,k;h)}) + \\
&\quad + s_1 p_{det}(i, N+1; h)\mathbb{E}(\prod_{l=i}^{N} \tilde{g}_l(s_1, s_2)^{n(i,l,N+1;h)}) \\
&= s_2 \sum_{k=i}^{N} p_{det}(i, k; h) \prod_{l=i}^{k} \mathbb{E}(\tilde{g}_l(s_1, s_2)^{n(i,l,k;h)}) + \\
&\quad + s_1 p_{det}(i, N+1; h) \prod_{l=i}^{N} \mathbb{E}(\tilde{g}_l(s_1, s_2)^{n(i,l,N+1;h)}) \\
&= s_2 \sum_{k=i}^{N} p_{det}(i, k; h) \prod_{l=i+1}^{k} \sum_{n=0}^{\infty} p(n; i, j, k; h)(\tilde{g}_l(s_1, s_2))^n + \\
&\quad + s_1 p_{det}(i, N+1; h) \prod_{l=i+1}^{N} (\sum_{n=0}^{\infty} p(n; i, l, N+1; h)(\tilde{g}_l(s_1, s_2))^n).
\end{aligned}$$

For interval N we have $\tilde{g}_N(s_1, s_2) = s_2 p_{det}(N, N) + s_1 p_{det}(N, N+1)$.

Now we can determine $\tilde{g}_0(s_1, s_2)$ point wise. Note that we only need to compute $\tilde{g}_0(s_1, s_2)$ in $m + 1$ points to give a good approximation of the first m derivatives of this function for a certain point. With this derivatives we are able to compute the first m moments of the size at time $t = T$ or to approximate $\mathbb{P}(X = n)$ for all $n \leq m$.

With this generating function we can also compute the probability of extinction and the expected final size of the epidemic. We can use the same reasoning as before. The progenies of all herds, infective at time $t = T$ have to go extinct. This occurs with probability q^{X_0}. So the probability of extinction is $\sum_{k=0}^{\infty} \mathbb{P}(X_0 = k)(q(N + 1))^k = \tilde{g}_0(q(N + 1), 1)$. Note that with these s_1 and s_2 the model is exactly the same as the model in Section 2.4.2. For the expected final size we add the expected number of infected herds, already detected, to the expected size of the progeny of all the infective herds at time $t = T$ (again including the herds infective at that time). This is $\frac{d}{ds_1} g_0(s_1, 1)|_{s_1=1} \bar{G}(N + 1) + \frac{d}{ds_2} g_0(1, s_2)|_{s_2=1}$. We can only use this property if after time $t = T$ the infection and detection rate are proportional to each other, because otherwise we need to know the ages of the infective herds at time $t = T$.

In a non-varying environment, we know for $q = 1$ the "speed" at which $\mathbb{P}(Z(t) > 0)$ decreases, for $t \to \infty$, where $Z(t)$ is the number of infective individuals at time t, descending from one individual infected at time $t = 0$ [39]. So we can give an upper bound for the probability the disease is already extinct at a certain time. We do this by assuming that all infective herds at time $t = T$ are infected at that time. So all herd at time T have age $\tau = 0$.

Now with the notation S_t for the probability that the progeny of a herd infected at time T, will not survive until time $T + t$ and with $Z(t)$ for the number of infected herds at time t, we have that:

$$\mathbb{P}(Z(T + t) = 0) = \sum_{k=0}^{\infty} \mathbb{P}(Z(T) = k)(S_t)^k = \tilde{g}(S_t, 1).$$

2.4.5 Discussion

1. By using this iterative method, we can compute the expected final size of an outbreak very fast, but for computing any higher moments of this final size, we need substantially more computational effort. By using simulations, it is possible to estimate these higher moments too. (see [49])

2. The lack of memory property is very important for our computations. We used it to write the formula of $q(i)$ in a 'convenient' form, with expectations in front of every $q(l)^n$.

   The speed of computations also heavily depends on the lack of memory property of the infection rate. In each interval the number of infections by one herd is Poisson distributed. We can easily integrate h out of the formula for $q(i)$, for Poisson distributed numbers of infections. (With the given distribution of H).

2.5 Results

In this section we use the Dutch classical swine fever epidemic of 1997 as example. We use the same data as Klinkenberg et al. in [49]. However Klinkenberg et al. used simulations and for every simulation the parameters were (pseudo-) randomly chosen from the distributions of the parameters, while we used only estimations of parameters for our computations.

We consider the following set of control measures:

A Total transport prohibition

B Killing of young piglets, in combination with a breeding ban

C Vaccination of all piglets (not sows) at multiplier herds, followed by recurrent vaccination of newborn piglets

D Single vaccination of all pigs at finishing herds

E Vaccination of piglets on arrival at finishing herds

One measure or a combination of measures is called a *control strategy* or *scenario*. The effects of different scenarios on the fraction of infective animals in a multiplier, in a finisher and the possibility of transport-infections ($\phi_m(t)$, $\phi_f(t)$ and $\phi_{tr}(t)$ respectively) are given in Table 2.1 (This table is adapted from [49]). For example, the 0 for ϕ_{tr} in strategy A means that infection is not possible by transport contacts. As another example, $\phi_f(t) = t/100$ for $t \leq 100$ in strategy D means that at time t a fraction $t/100$ of the animals in a finisher is infective. Note that the first 12 strategies are in a constant environment. Therefore, for those strategies we can compute properties of the spread of the disease directly.

The probabilities of extinction for various strategies are given in Table 2.2. We also compute the expected final size of an epidemic (Table 2.3) We only need to compute this for the strategies with almost sure extinction. The expected number of infected multipliers, while the initially infected herd was a finisher is denoted by G_{fm}. In a similar way we define G_{mm}, G_{mf}, G_{ff}. The last column of Table 2.3 is the expected number of infected herds, when initially five multipliers and five finishers were infected. Note that we cannot compare the expected final size with the results of Klinkenberg et al. (Table 6 in [49]), because they use the median of 1000 simulations, not the mean. The results of Klinkenberg et al. are given in Table 2.4. The 95% confidence intervals of the final size are also taken from [49] and were estimated by simulation.

We used the generating function for the number of infected and infective herds at time $t = T$, the point in time after which the parameter values are assumed to be constant, to estimate the size at time T (which varies for different strategies). We estimate the first two moments of the size at time T for the four pure strategies in a varying environment B, C, D, E. (Table 2.5) Especially the results of B and C are of interest because they will give an idea of the variance of the final size of a herd.

On the next pages we give covariance matrices (for different strategies and different initially infected herds) of the number of infected multipliers not yet detected, the number of infected multipliers already detected, the number of infected finishers not yet detected and the number of finishers already detected, respectively. We denote by $Var_i(J)$ the covariance matrix for an initially infected herd of type i, while the strategy is J. By comparing the values in these matrices to the results in Table 2.5, we know how much trust we may put in the expected sizes.

Up to now we used exact parameters. In reality we usually cannot estimate the parameters exactly. In order to get some insight into how the computed properties depend on the values of the different parameters, we varied one parameter while keeping the other parameters constant. The results for the strategies B and D are given in Figures 2.1 and 2.2. The scales on the axes differ for the different varying parameters. On the x-axes we put the ratio between the value of the parameter and the point estimator of that parameter. We see that for these strategies, the parameters r and R have little influence on the computed quantities, while α and β do have significant influence.

$$Var_m(B) = \begin{pmatrix} 0.85 & 2.19 & 0.52 & 8.43 \\ 2.19 & 16.97 & 2.63 & 56.60 \\ 0.52 & 2.63 & 1.06 & 10.21 \\ 8.43 & 56.60 & 10.21 & 237.42 \end{pmatrix}$$

$$Var_f(B) = \begin{pmatrix} 0.38 & 0.98 & 0.23 & 3.43 \\ 0.98 & 8.02 & 1.17 & 23.84 \\ 0.23 & 1.17 & 0.48 & 4.13 \\ 3.43 & 23.84 & 4.13 & 90.10 \end{pmatrix}$$

$$Var_m(C) = \begin{pmatrix} 0.92 & 0.66 & 0.20 & 0.68 \\ 0.66 & 1.68 & 0.25 & 1.00 \\ 0.20 & 0.25 & 0.21 & 0.26 \\ 0.68 & 1.00 & 0.26 & 1.87 \end{pmatrix}$$

$$Var_f(C) = \begin{pmatrix} 0.92 & 0.67 & 0.20 & 0.67 \\ 0.67 & 1.67 & 0.25 & 1.00 \\ 0.20 & 0.25 & 0.21 & 0.26 \\ 0.67 & 1.00 & 0.26 & 1.87 \end{pmatrix}$$

$$Var_m(D) = \begin{pmatrix} 56.4 & 27.5 & 97.8 & 48.7 \\ 27.5 & 18.4 & 53.2 & 24.2 \\ 97.8 & 53.2 & 197.9 & 90.6 \\ 48.7 & 24.2 & 90.6 & 54.5 \end{pmatrix}$$

$$Var_f(D) = \begin{pmatrix} 23.0 & 11.4 & 39.5 & 17.6 \\ 11.4 & 8.1 & 22.2 & 10.5 \\ 39.5 & 22.2 & 80.9 & 32.5 \\ 17.6 & 10.5 & 32.5 & 18.0 \end{pmatrix}$$

$$Var_m(E) = \begin{pmatrix} 55.0 & 29.0 & 65.3 & 46.8 \\ 29.0 & 21.0 & 38.9 & 29.7 \\ 65.3 & 38.9 & 95.5 & 59.4 \\ 46.8 & 29.7 & 59.4 & 53.0 \end{pmatrix}$$

$$Var_f(E) = \begin{pmatrix} 24.3 & 12.9 & 28.3 & 18.3 \\ 12.9 & 9.9 & 17.2 & 12.0 \\ 28.3 & 17.2 & 41.9 & 22.7 \\ 18.3 & 12.0 & 22.7 & 19.0 \end{pmatrix}$$

In the constant environment the only thing that matters is the ratio $\frac{\beta}{\alpha}$. For varying environments we varied α, while keeping $\frac{\beta}{\alpha}$ and $\frac{\beta_{tr}}{\alpha}$ constant. The results are varying, but not very much.

2.5.1 Discussion

1. We cannot estimate the parameters exactly (see Chapter 3). Because the expected final size, the probability of extinction and the generating function for the number of infective herds at a certain moment heavily depend on some of the parameters, we cannot really use the computed values in a quantitative way. But we can use them to *compare* different scenarios, while using the same parameter values.

2. We computed the probability of extinction for the progeny of one infective herd, but in reality we may have more than one infected herd at the moment the first measures are implemented, so the probability of extinction may be much less than the computed q. We assumed that different infected herds infect other herds independently of each other. If we simplify the model by assuming that all herds at time t=0 have age 0, we can estimate the expected final size by the number of initially infected herds times the estimated final size of one infected herd. In the same way we can estimate the probability of extinction by the computed probability for one herd, to the power of the number of initially infected herds. From previous outbreaks of FMD, CSF and Avian Influenza in the Netherlands one can see that as a rule several farms are already infected at the moment the first case is suspected or confirmed.

3. The results are very much the same as the results of Klinkenberg et al. but we do not need simulations to estimate the properties. Also, our method is much faster than using simulations. In Table 6 of [49] it is suggested that the expected final size, with initially five infected multipliers and five infected finishers, is estimated by simulations. In reality the median of 1000 simulations is used in that paper. This explains why our results differ on that point.

4. Note that strategy ABC gives an extremely high expected final size. Klinkenberg et al.[49] did not even give this final size. This is because for the given parameters the process is very near to a process with a positive probability of surviving i.e. a positive probability of an infinite final size.

5. Using the expected size at $t = T$ and the variance of this size, we see that the expected final size is not very informative in some scenarios. Large variance may imply that a large set of "numbers of infected herds" have a not-ignorable probability, even if we know the exact parameters.

6. The computed covariance matrices do have values of different order in them. Consider strategy B. The variance of the number of finishers and multipliers already detected is much higher than the variance of the number of finishers and multipliers that are still alive at time $t = T$. This is because the infection rate decreases for this strategy. In the start of the process one infective herd may infect several other herds with a relatively high probability, while after some time infections become rare. This is why the expected number of infective herds at $t = T$ is much less than the expected number of herds infected during the epidemic.

7. For constant environments we are interested in the variance of the final size. Note that in all of the strategies with almost sure extinction in a constant environment, only one type of herd can be infected, so we only have to deal with one type of herd. We use the underlying Galton-Watson process to deduce the variance of the final size (see Section 2.11 of [39]). If we take m for the expected number of direct infections by one infective herd, the expected final size is $\frac{1}{1-m}$ and the variance is $\frac{m(1+m)}{(1-m)^2}$. Note that for a large expected final size, the variance will be of the square order of the final size. Due to this large variance, we cannot give exact quantitative predictions about the final size of an epidemic. This also indicates that there is a large intrinsic uncertainty in the problem.

8. If the infection rates are increasing in time, the expected final size will decrease for increasing r, while for decreasing infection rates the expected final size will increase for increasing r. This is because r can be seen as a parameter describing the speed of the process in a constant environment. Large values of r correspond to short generation lengths compared to small r. So, if the infection rate is increasing in time, the n-th generation for small values of r will have a larger infection rate than the n-th generation for large r.

   The effect of decreasing R, the parameter describing the random effects at the start of the within-herd spread, is less clear from formulae and definitions. From the figures we can only see that R has relative small effect on the expected final size and the probability of extinction.

9. Some parameters heavily influence the expected final size and the probability of extinction. We have already seen that in a constant environment r and R do not influence these quantities. For predictions the ratios $\beta : \beta_{tr} : \alpha$ are most important. So the estimation effort is best devoted towards estimating these ratios.

2.6 Conclusion and final remarks

By using a stochastic model, we could estimate the probability of extinction and the expectation of the final size of an epidemic in a varying environment. We only need that the environment is not varying anymore after some given time. We use a branching process in varying environments, depending on the age of the individuals. In the constant environment the probability of extinction and the expected final size are known. By using an iterative process, we computed this probability and size for the time the environment is still varying.

By using generating functions, we can compute many important properties, like the moments of the final size, and a lower bound for the probability that the process has gone extinct after a given time. These generating functions are very useful especially if in a constant environment the expected number of infections after a certain age does not depend on that age. This only holds if the infection and detection rate are in direct proportion, for the non-varying environment.

It is difficult to estimate the parameters for the computations. Some of the properties computed in this model, like the probability of extinction heavily depend on the parameters α and β, describing the infection and detection rate. The model presented in this chapter may be useful to *compare* the effect of different measures, but it is very dangerous to use this model for absolute quantitative predictions.

We used the model to describe the spread of classical swine fever. It is worth investigating if the same model, with other parameters, may be used to describe the spread of other animal diseases, with culling at detection, or even human diseases, which lead to strict quarantine of detected infected individuals (and suspected cases) and contact and movement restrictions. Emerging infections such as SARS are possible examples of this.

[1]Table 2.1 is reprinted from *Mathematical Biosciences*, 186 (2), Klinkenberg, D.; Everts-van der Wind, A.; Graat, E.A.M. and De Jong, M.C.M., Quantification of the effect of control strategies on classical swine fever epidemics, 145-173, Copyright (2003), with permission from Elsevier.

Table 2.1: The effect of different strategies on $\phi_f(t)$, $\phi_m(t)$ and $\phi_{tr}(t)$.[1]

Strategy	Time interval	$\phi_f(t)$	$\phi_m(t)$	$\phi_{tr}(t)$
none	$t > 0$	1	1	1
ABCD	$t > 0$	0	365/1975	0
ABC	$t > 0$	1	365/1975	0
ABD	$t > 0$	0	1	0
AB	$t > 0$	1	1	0
ACD	$t > 0$	0	1009/1975	0
AC	$t > 0$	1	1009/1975	0
AD	$t > 0$	0	1	0
A	$t > 0$	1	1	0
BCD	$t > 0$	0	365/1975	0
CD	$t > 0$	0	1009/1975	0
DE	$t > 0$	0	1	1
BC	$0 < t \leq 100$	$1 - t/100$	365/1975	0
	$t > 100$	0	365/1975	0
BDE	$0 < t \leq 70$	0	$1 - 23t/1975$	1
	$t > 70$	0	365/1975	0
BD	$0 < t \leq 70$	$t/100$	$1 - 23t/1975$	1
	$70 < t \leq 100$	0.7	365/1975	0
	$100 < t \leq 170$	$1.7 - t/100$	365/1975	0
	$t > 170$	0	365/1975	0
BE	$0 < t \leq 70$	$1 - t/100$	$1 - 23t/1975$	1
	$70 < t \leq 100$	$1 - t/100$	365/1975	0
	$t > 100$	0	365/1975	0
B	$0 < t \leq 70$	1	$1 - 23t/1975$	1
	$70 < t \leq 170$	$1.7 - t/100$	365/1975	0
	$t > 170$	0	365/1975	0
C	$0 < t \leq 100$	$1 - t/100$	1009/1975	0
	$t > 100$	0	1009/1975	0
D	$0 < t \leq 100$	$t/100$	1	1
	$t > 100$	1	1	1
E	$0 < t \leq 100$	$1 - t/100$	1	1
	$t > 100$	0	1	1

Table 2.2: Probability of extinction for various strategies.

Strategy	$q_m(0)$	$q_f(0)$
none	0.32	0.51
ABCD, ABC, ABD	1	1
ACD, AD, BCD, CD		
AB	0.61	0.61
AC	0.80	0.80
A	0.61	0.61
DE	0.69	0.46
BC, BDE, BD, BE, B, C	1	1
D	0.38	0.61
E	0.39	0.59

Table 2.3: Expected final size for various strategies.

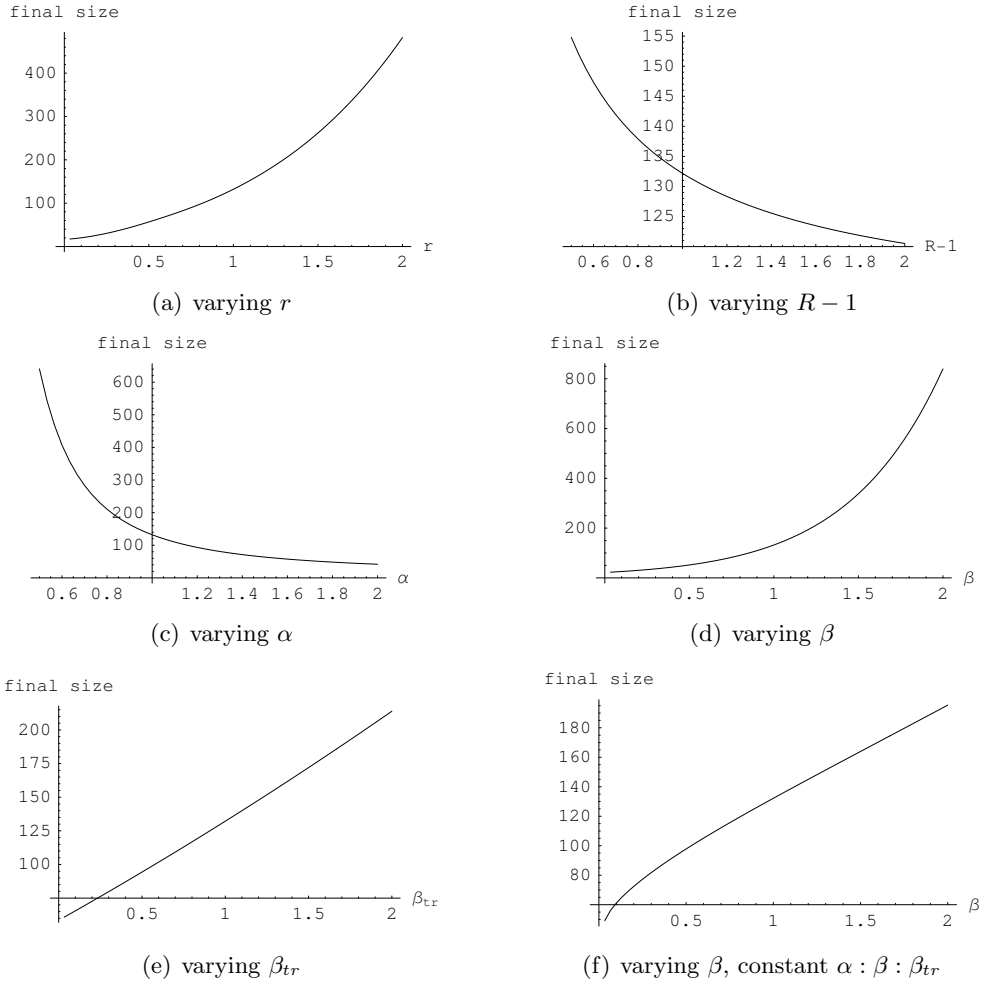
Strategy	G_{mm}	G_{mf}	G_{fm}	G_{ff}	$G_{5m+5f;f+m}$
ABCD	1.2	0	0.2	1	11.8
ABC	7.5	35.4	6.5	36.4	429.4
ABD	5.7	0	4.7	1	56.9
ACD	1.7	0	0.7	1	17.3
AD	5.7	0	4.7	1	56.9
BCD	1.2	0	0.2	1	11.8
CD	1.7	0	0.7	1	17.3
BC	1.1	0.8	0.3	1.8	21.6
BDE	2.7	3.6	0.9	2.0	45.4
BD	3.6	8.5	1.3	4.0	87.2
BE	3.2	5.8	1.2	3.2	66.6
B	4.5	13.2	1.9	6.8	132.3
C	2.4	1.0	1.4	2.0	33.8

Table 2.4: Median and 95% confidence interval of the final size.

Strategy	median of final size	95% CI
ABCD	11	$10-16$
ABD	42	$13-452$
ACD	16	$10-35$
AD	42	$13-452$
BCD	11	$10-16$
CD	16	$10-35$
BC	21	$12-38$
BDE	35	$15-69$
BD	67	$20-157$
BE	53	$22-110$
B	108.5	$37-238$
C	33	$15-73$

Table 2.5: Expected sizes at time T.

Strat.	initially inf.	undet. m	detected m	undet. f	detected f
B	multiplier	0.39	3.98	0.47	12.57
B	finisher	0.18	1.61	0.22	6.50
C	multiplier	0.35	1.72	0.14	0.82
C	finisher	0.35	0.72	0.14	1.82
D	multiplier	5.26	4.24	9.93	6.12
D	finisher	2.39	1.72	4.57	3.28
E	multiplier	5.34	4.55	6.88	6.27
E	finisher	2.64	2.00	3.45	3.65



(a) varying r

(b) varying $R-1$

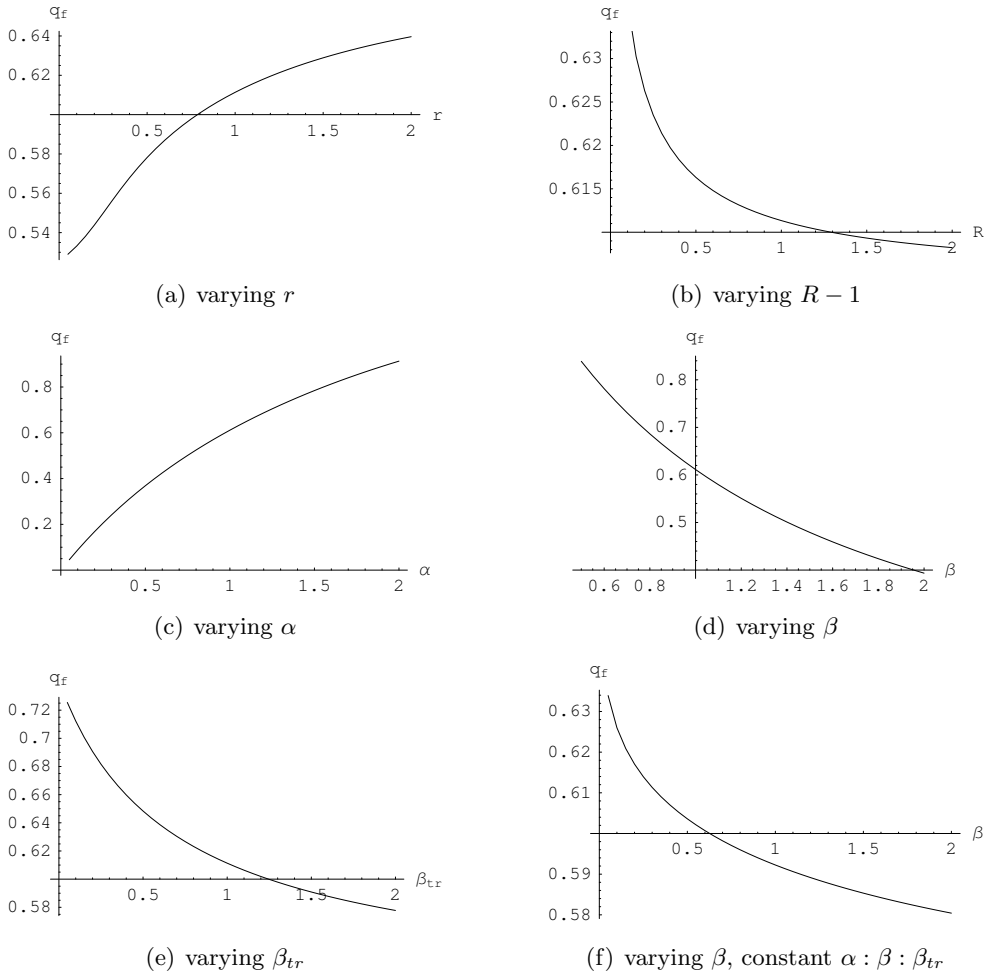
(c) varying α

(d) varying β

(e) varying β_{tr}

(f) varying β, constant $\alpha : \beta : \beta_{tr}$

Figure 2.1: The expected final size of an epidemic of classical swine fever initiated by 5 multipliers and 5 finishers, with control measure B. The parameters at the x-axis are scaled, such that the point estimator used in this chapter, has value 1.



(a) varying r

(b) varying $R-1$

(c) varying α

(d) varying β

(e) varying β_{tr}

(f) varying β, constant $\alpha : \beta : \beta_{tr}$

Figure 2.2: The probability of extinction of an epidemic of classical swine fever initiated by 1 finishers, with control measure D. The parameters at the x-axis are scaled, such that the point estimator used in this chapter, has value 1.

Chapter 3

Estimation in branching processes with restricted observations

3.1 Introduction and motivation

It is known that it is impossible to consistently estimate more than two moments of the offspring-distribution in a supercritical Galton-Watson process if only the generation sizes X_n of the process are observed, (Theorem 1.3 of [34]). However, it is not a-priori clear what one can estimate consistently in a situation, where in generation n, any individual is detected with unknown probability π and the numbers of these detected individuals are the only observations we have. The detected individuals may produce a reduced offspring.

The kind of partial observations we are dealing with is especially interesting for estimation in epidemics of infectious diseases. If the number of susceptible individuals (where an individual may also refer to a herd instead of an individual animal) is very large, we may describe the start of an idealized epidemic by a Galton-Watson process, where discrete points in time index the generations (see e.g. [3] or Chapter 2). As soon as an infectious disease is observed in an individual, it stops being infective because of isolation (in the case of human infections) or culling (in the case of very contagious animal diseases like classical swine fever, foot and mouth disease or avian influenza).

In the time interval of detection, the individual was only infectious during a fraction of the interval length which implies that observed individuals have reduced offspring. In this epidemiological setting, an individual that is infectious but not detected in a certain generation, will still be infectious in the next one, so a surviving individual will cause at least one infective individual in the next generation, namely itself. Our task is to estimate parameters of the offspring distribution using only this partial information. This interpretation should be compared to the work in [12]. In that paper, estimation takes place under the assumption that one also observes, in addition to individuals without further offspring, the total number of infectious individuals at the beginning, and at the end of the observation period.

If X_n denotes the generation sizes of a branching process and π_n a known sequence converging to π, Jacob and Peccoud [38] have shown that if the number of observations in generation $n+1$ is binomially distributed with parameters X_n and π_n and the offspring distribution has a finite fourth moment, then it is also possible to estimate the first two moments of the offspring distribution consistently on the explosion set (i.e. the set where $\lim_{n\to\infty} X_n = \infty$). In their assumptions the observed individuals may produce offspring, but the offspring of these individuals is supposed to be distributed like the offspring of L unobserved individuals, where L is a non-negative integer.

Our set-up differs in two aspects from [38]. First, we are interested in the case where $\pi_n = \pi$ is constant but unknown. Furthermore, we assume that the offspring distribution of unobserved and observed individuals have a finite fourth moment, but no further assumptions are made about the offspring distributions. Our methods are also quite different; our martingales are based on observable quantities. Besides proving some results analogous to [38], we also show that π can under certain circumstances be estimated consistently.

Our main interest is in estimating the offspring mean and the parameter π, because those parameters are extremely important for decisions about measures to be taken to stop an epidemic. We are able to estimate the offspring mean very efficiently. On the explosion set, we are (under certain conditions) able to consistently estimate two other functions of π and the parameters of the offspring distribution. These three estimators will lead to a system of three equations. For many models we also have three unknowns, namely the offspring mean, the offspring variance and π. In principle, we can therefore often estimate these quantities. However, it turns out that what is theoretically

possible, is not always practically feasible due to extreme slow convergence of the second and third estimator. Note that one can only hope to obtain consistency if the process explodes, otherwise the number of observations and the number of involved individuals will be finite.

In the next section we set things up formally. In Section 3.3 we give consistent estimators for three functions of the parameters and results about the rate of convergence of these estimators. We apply this to real data from the 1997 epidemic of classical swine fever in The Netherlands. In Section 3.4 we prove the consistency of two estimators for the offspring mean on the explosion set of the branching process. In Section 3.5 we estimate a second function of the parameters. This function can be interpreted as a second moment, as we will explain. In Section 3.6 we show that the estimator for a third function of parameters is consistent.

3.2 Formal set-up

We let G_n be the collection of infected individuals at the discrete time instants n, $n = 0, 1, 2, \ldots$, and we let $X_n = |G_n|$ denote its cardinality. In our context, the X_n are *not* observable. The dynamics from time n to time $n + 1$ is as follows.

Between time n and $n+1$, a certain (random) number of the infected individuals in G_n is detected; we assume that each infected individual is detected with probability π during this time interval, independently of each other. The parameter π is unknown. The collection of detected individuals between time n and time $n + 1$ is denoted by D_{n+1} and the number of individuals in D_{n+1} is denoted by Z_{n+1}; this random quantity is observable. So, given X_n, Z_{n+1} has a binomial distribution with parameters X_n and π. Individuals in $\cup_n D_n$ produce no offspring.

An individual in G_n which is detected (and which therefore produces an element in D_{n+1}) may also produce offspring in G_{n+1} (in our terminology, 'offspring' always means 'direct offspring'). An individual in G_n which is not detected will remain infective, and possibly infect other individuals. The offspring of such an individual in G_{n+1} therefore consists of at least one individual, namely itself. Note that as a result, one physical individual corresponds to various individuals of the process. The whole process now constitutes a two-type branching process.

The offspring distributions of detected and undetected individuals are different. We denote by m_+ the expected number of X-offspring of an infected individual at time n (that is, offspring in G_{n+1}), given that it is not detected between time n and time $n+1$. Similarly, m_- is the expected number of X-offspring of an infected individual given that it is detected. The corresponding variances are denoted by σ_+^2 and σ_-^2 respectively. In formulas, this reads as follows:

$$\begin{aligned} m_+ &:= \mathbb{E}(X_1|X_0=1,Z_1=0), \\ m_- &:= \mathbb{E}(X_1|X_0=1,Z_1=1), \\ \sigma_+^2 &:= \mathbb{E}((X_1-(m_+))^2|X_0=1,Z_1=0), \\ \sigma_-^2 &:= \mathbb{E}((X_1-(m_-))^2|X_0=1,Z_1=1). \end{aligned}$$

Finally, we write m for the unconditional expected number of X-offspring of an infected individual:

$$m = \mathbb{E}(X_1|X_0=1) = (1-\pi)m_+ + \pi m_-.$$

Similarly, the unconditional variance is denoted by σ^2. We have from Lemma 2.1 of [34] that for a random variable Y, an event F and F^c the complement of F,

$$\begin{aligned} Var(Y) &= \mathbb{P}(F)Var(Y|F) + \mathbb{P}(F^c)Var(Y|F^c) + \\ &\quad +(\mathbb{E}(Y|F) - \mathbb{E}(Y|F^c))^2\mathbb{P}(F)\mathbb{P}(F^c). \end{aligned}$$

Applying this with $F = \{Z_1 = 0\}$ yields

$$\sigma^2 = Var(X_1|X_0=1) = (1-\pi)\sigma_+^2 + \pi\sigma_-^2 + (m_+ - m_-)^2\pi(1-\pi).$$

We assume that the offspring distributions have finite fourth moment, i.e. for $i \in \{0,1\}$

$$\mathbb{E}([X_1 - \mathbb{E}(X_1|X_0=1,Z_1=i)]^4|X_0=1,Z_1=i) < \infty.$$

Define A as the explosion set, that is, the set where $\lim_{n\to\infty} X_n = \infty$. Because A is a tail-event, conditioning on this event is strictly speaking not proper for estimation purposes, but analysing the behaviour of the process on the set A is necessary, for only on this set we obtain infinitely many observations.

We are able to estimate several functions of m_+, m_-, σ^2 and π on A. One of those functions is m, the other two are given by

$$\begin{aligned} \gamma := \gamma(m_+, m_-, \sigma^2, \pi) &:= (1-\pi)m + \pi\sigma^2 + \pi m^2 + m^2 - 2\pi m m_-, \\ \gamma_* := \gamma_*(m_+, m_-, \sigma^2, \pi) &:= (m^2+m)\gamma - 2m^3 + 2\pi m^2 m_-. \end{aligned}$$

The reason for these somewhat complicated expressions will become clear soon.

3.3 Results and application

3.3.1 The main result

In this section, consistent estimators for three different functions of the parameters are given.

Theorem 3.3.1 *Using the notation and assumptions of Section 2, we have as* $n \to \infty$

$$\begin{array}{llcll} (a) & \bar{m}_n & \to & m, & \text{a.s. on } A, \\ (b) & \tilde{m}_n & \to & m, & \text{a.s. on } A, \\ (c) & n^{-1}\tilde{S}_n(\tilde{m}_n) & \to & \gamma, & \text{in probability on } A, \\ (d) & n^{-1}\tilde{S}^*_n(\tilde{m}_n) & \to & \gamma_*, & \text{in probability on } A, \end{array}$$

where

$$\begin{aligned} \bar{m}_n &= \frac{Z_{n+1}}{Z_n}, \\ \tilde{m}_n &= \frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} Z_i}, \\ \tilde{S}_n(m) &= \sum_{i=1}^{n} (Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - m\right)^2, \\ \tilde{S}^*_n(m) &= \sum_{i=1}^{n} (Z_i+1)\left(\frac{Z_{i+2}}{Z_i+1} - m^2\right)^2. \end{aligned}$$

In principle, this theorem gives three equations with four unknowns, namely m, π, σ^2 and m_-. If we have further information, or make further assumptions about the relation between m and m_-, then we can estimate all parameters consistently, in theory at least.

The speed of convergence of our estimators is given by the following theorems.

Theorem 3.3.2 *The random variables*

$$\Big(\sum_{i=1}^{n} Z_i\Big)^{\frac{1}{2}}(\tilde{m}_n - m)$$

converge in distribution to a sum of three normal random variables with zero mean and finite variance.

This theorem implies that $(\tilde{m}_n - m)$ converges with a rate of order $\Big(\sum_{i=1}^{n} Z_i\Big)^{-\frac{1}{2}}$ to 0.

Theorem 3.3.3 *As $n \to \infty$, for all $\delta > 0$ we have*

$$\begin{aligned} (a) \quad & n^{\frac{1}{2}-\delta}\left(n^{-1}\tilde{S}_n(\tilde{m}_n) - \gamma\right) \;\to\; 0, \\ (b) \quad & n^{\frac{1}{2}-\delta}\left(n^{-1}\tilde{S}^*_n(\tilde{m}_n) - \gamma_*\right) \;\to\; 0, \end{aligned}$$

in probability on A.

This theorem implies that $n^{-1}\tilde{S}_n(\tilde{m}_n) - \gamma$ and $n^{-1}\tilde{S}^*_n(\tilde{m}_n) - \gamma_*$ converge with a rate of order at least $n^{-\frac{1}{2}+\delta}$, for any $\delta > 0$.

The proofs of Theorem 3.3.1(a) and (b) and Theorem 3.3.2 are given in Section 3.4. The proofs of Theorem 3.3.1(c) and Theorem 3.3.3(a) are given in Section 3.5. Finally, the proof of Theorem 3.3.1(d) is given in Section 3.6. Theorem 3.3.3(b) can be proved in exactly the same way as Theorem 3.3.3(a) and the proof is omitted.

3.3.2 Application to an epidemic model

In this subsection, we apply our results to a concrete example from epidemic theory. We analyse the discrete approximation of the standard stochastic SIR epidemic (see e.g. [27] and Chapter 1, Section 1.3.2), where we take the number of susceptible individuals to be infinite.

The Model:

We assume that if an infective individual is not detected in a certain interval, the number of new infections by this infective individual is Poisson distributed

with parameter λ. Since an individual that is not detected remains infective itself, this leads to

$$\begin{aligned} m_+ &= \lambda + 1, \\ \sigma_+^2 &= \lambda. \end{aligned}$$

Next we need to make a choice for what happens during the interval that an individual is detected. In order to keep the model rather general we assume that the detected individual was infective during a (known) fraction ϕ of the detection interval. It then follows that

$$\begin{aligned} m_- &= \phi\lambda, \\ \sigma_-^2 &= \phi\lambda, \end{aligned}$$

and hence

$$\begin{aligned} m &= (1-\pi)(\lambda+1) + \phi\lambda\pi, \\ \sigma^2 &= (1-\pi)\lambda + \phi\lambda\pi + \pi(1-\pi)((1-\phi)\lambda+1)^2. \end{aligned}$$

One remark: if we assume that the detection time is uniformly distributed over the interval of detection, m_- will be $\lambda/2$. The variance of the offspring in the interval of detection will be slightly larger than $\lambda/2$ because the randomness of the detection time will cause some extra variance. In fact the variance would be $\lambda/2 + \lambda^2/12$.

We are particularly interested in estimating m (which describes the mean growth of the number of infectious individuals) and π (needed to estimate the number of infectious individuals at a certain time, which is very important in order to make decisions about measures to stop the epidemic).

In the context of the present example, Theorem 3.3.1 gives that on A,

$$\begin{aligned} \tilde{m}_n &\to (1-\pi)(\lambda+1) + \phi\lambda\pi && \text{a.s.}, \\ \tfrac{1}{n}\tilde{S}_n(\tilde{m_n}) &\to (1-\pi)\lambda + (1-\pi)^2 + (1-\pi)(\lambda+1)^2 + \phi\lambda\pi && \text{in probability}, \\ \tfrac{1}{n}\tilde{S}_n^*(\tilde{m_n}) &\to (m^2+m)\gamma + 2m^2(1-\pi)(\lambda+1) && \text{in probability}. \end{aligned} \tag{3.1}$$

For ease of notation, we did not expand the last quantity at the right-hand side. In (3.1) we have, after substituting the estimates for m, γ and γ_*, three equations with two unknowns, λ and π. At this point, it seems that the third

equation does not help us very much. There are several ways to proceed now.

First of all, we can ignore the third equation and solve the other two for λ and π. However, it turns out that $n^{-1}\tilde{S}_n(\tilde{m}_n)$ converges very slow, and we need a huge number of generations to obtain reliable estimates (see the next subsection). But there is a way to use the information contained in the third equation in a meaningful way. To this end, we reparametrise the epidemic process by using m and π instead of λ and π. The parameter m is estimated by $\tilde{m}_n$ while we may use the combination $(1+\tilde{m}_n^{-1})n^{-1}\tilde{S}_n(\tilde{m_n}) - \tilde{m}_n^{-2}n^{-1}\tilde{S}_n^*(\tilde{m_n})$ to estimate $2(1-\pi)(\lambda+1)$. Since $m = (1-\pi)(\lambda+1) + \phi\lambda\pi$ we may write:

$$\lambda + 1 = \frac{m + \phi\pi}{(1-\pi) + \phi\pi},$$

so again we have a system of two equations with two unknowns. From simulation results it turns out that we can give reasonable estimates for π much faster than in the case where we use only the estimators for m and γ, but for practical purposes our new estimators for π still converges too slow.

The Data:

We did the analysis of the data of the 1997 Dutch classical swine fever (CSF) outbreak as treated in [56]. In that paper the outbreak is modelled by a Galton-Watson process. Because of changing measures of the government, the parameters π and λ differ for different stages of the epidemic. For this reason the epidemic is divided into 5 stages. The time unit is one week and it is assumed that on average detections take place in the middle of a time interval,

Table 3.1: Estimates from the 1997 Dutch CSF-outbreak; The epidemic is divided into 5 stages (st.), 4 of which are used. The number of weeks (wk.) and the number of detected farms (obs.) during the stages are given. The values obtained by computing m, γ and γ_* using the MLE's of π and λ as given in [56] are compared with our current estimates.

			mle's		computed			estimates		
st.	wk.	obs.	π	λ	m	γ	γ_*	m	γ	γ_*
2	10	101	0.4	0.6	1.08	2.376	3.098	1.15	1.03	0.942
3	8	160	0.5	0.7	1.025	2.220	2.822	1.04	1.87	2.57
4	9	107	0.3	0.3	0.955	1.928	1.940	0.857	0.813	0.935
5	30	51	0.05	0.25	1.194	2.631	3.505	0.880	0.859	0.867

Table 3.2: Estimated values at different generations of simulated data when $\lambda = 0.6$, $\pi = 0.4$ and $\phi = 0.5$.

generation (n)	$\tilde{m}_n$	$n^{-1}\tilde{S}_n(\tilde{m}_n)$	$(n-1)^{-1}\tilde{S}_{n-1}(\tilde{m}_n)$	estimated π
20	1.112	2.907	2.745	–
50	1.082	2.849	3.142	–
100	1.082	2.928	3.750	–
250	1.080	2.478	3.424	0.4601
500	1.080	2.522	3.288	0.2852
750	1.080	2.398	3.059	0.3318

so ϕ is set to 0.5. In [56], an expensive algorithm is used find the maximum likelihood estimators for π and λ in the different stages. Klinkenberg et al.([48], Chapter 6) already showed that these maximum likelihood estimates are not very reliable.

We will use our estimators to estimate m and γ for the different stages. We omit the first stage, because for that stage we have only one observation. We compare our estimates with the m and γ computed from the maximum likelihood estimates of λ and π given in [56]. The results are given in Table 3.1. We also give the duration of the stage (in weeks) and the number of observed individuals in a stage of the epidemic in this table.

In the second, third and fourth stage of the epidemic our estimate for m seems to be rather good, as we might expect. The estimate for γ does not seem to be very informative. In the final stage of the epidemic only few cases were observed and there were many weeks without any observation. Due to this few observations we may expect our estimators not to converge very fast. In none of the stages we could estimate π and λ by using our estimated m and γ in (3.1), as the solutions of this system of equations gave no real π between 0 and 1. We have simulated the epidemic with the MLE's from [56] as the real parameter values to get some idea about the speed of convergence of the estimator of π (Table 3.2), we see that even after 750 weeks π is not accurately estimated.

3.4 Estimating the offspring mean m

In this section we discuss the two consistent estimators for m given in 3.3.1. We start with the estimator

$$\bar{m}_n := \frac{Z_{n+1}}{Z_n},$$

and Theorem 3.3.1(a) states that this estimator is indeed consistent on the explosion set A.

Proof of Theorem 3.3.1(a). The proof is based on a simple martingale argument. Let

$$M_n := \prod_{i=0}^{n} \frac{Z_{i+1}+1}{\pi X_i + 1}$$

Note that this is a (positive) martingale with respect to $\mathcal{F}_n$, the σ-algebra generated by $\{Z_{i+1}, X_i; 0 \le i \le n\}$. Since $\sup_n \mathbb{E}(M_n) \le 1$, the martingale convergence theorem implies that M_n converges almost surely to an almost surely finite random variable M.

We also need to show that M is strictly positive on A. To do this, we define

$$\bar{M}_n := \prod_{i=0}^{n} \frac{\pi(X_i+1)}{Z_{i+1}+1}.$$

Elementary computations yield

$$\begin{aligned} \mathbb{E}\Big[\frac{1}{Z_1+1}|X_0=k\Big] &= \frac{1}{\pi(k+1)}(1-(1-\pi)^{k+1}) \\ &\le \frac{1}{\pi(k+1)}, \end{aligned} \tag{3.2}$$

so $\bar{M}_n$ is a supermartingale with respect to $\mathcal{F}_n$. By the martingale convergence theorem we know that $\bar{M}_n$ converges almost surely to an almost surely finite random variable $\bar{M}$. Now write

$$M_n\bar{M}_n = \prod_{i=0}^{n} \frac{\pi(X_i+1)}{\pi X_i+1} = \prod_{i=0}^{n}\Big(1-\frac{1-\pi}{\pi X_i+1}\Big).$$

The X_i's almost surely grow exponentially on A, so

$$\sum_{i=1}^{\infty} \frac{1-\pi}{\pi X_i + 1} < \infty$$

almost surely on A, which implies that

$$M\bar{M} = \lim_{n\to\infty} M_n \bar{M}_n = \prod_{i=0}^{\infty} \Big(1 - \frac{1-\pi}{\pi X_i + 1}\Big) > 0, \tag{3.3}$$

almost surely on A. Because $\bar{M}$ is almost surely finite, (3.3) is only possible if M is almost surely positive on A.

Now since $X_{i+1}/X_i \to m$ a.s. on A (Theorem 2.1 of [34]), we have that $(X_{n+1} + \pi^{-1})/(X_n + \pi^{-1}) \to m$ a.s. on A. So we have

$$\frac{Z_{n+1}+1}{Z_n+1} = \frac{M_n/M_{n-1}}{M_{n-1}/M_{n-2}} \frac{\pi X_n + 1}{\pi X_{n-1} + 1} \to m, \qquad \text{a.s. on } A, \tag{3.4}$$

because from $0 < M < \infty$ a.s. on A, it follows that

$$\frac{M_n}{M_{n-1}} \to \frac{M}{M} = 1, \qquad \text{a.s. on } A.$$

On A, Z_n will almost surely tend to infinity, so $(Z_{n+1}+1)/(Z_n+1)$ will have the same limit as $(Z_{n+1})/Z_n$, proving the theorem.

Clearly, $\bar{m}_n$ does not use all the available information. To this end, we also considered the second estimator, namely

$$\tilde{m}_n := \frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} Z_i}.$$

Theorem 3.3.1(b) states that $\tilde{m}_n$ also is a consistent estimator for m. To prove this theorem, we start with a lemma from [35].

Lemma 3.4.1 *Let* $(S_n = \sum_{i=1}^{n} \xi_i, \mathcal{F}_n, n \geq 1)$ *be a martingale and let* $(U_n, n \geq 1)$ *be a non decreasing sequence of positive random variables such that* U_n *is* $\mathcal{F}_{n-1}$*-measurable. Then* $U_n^{-1} S_n \to 0$ *a.s. on the set*

$$\left\{ \lim U_n = \infty, \sum_{i=1}^{\infty} U_i^{-2} \mathbb{E}(\xi_i^2 | \mathcal{F}_{i-1}) < \infty \right\}.$$

Corollary 3.4.2 *On the explosion set A we have a.s. as $n \to \infty$,*

$$\frac{\sum_{i=1}^{n} Z_i}{\sum_{i=0}^{n-1} X_i} \to \pi.$$

This corollary is intuitively obvious because the denominator is the total number of individuals in the first n generations (including generation 0), while the numerator is the number of observed individuals in these generations. Here is a formal proof.

Proof of Corollary 3.4.2. Write

$$\frac{\sum_{i=1}^{n} Z_i}{\sum_{i=0}^{n-1} X_i} = \frac{\sum_{i=1}^{n} (Z_i - \pi X_{i-1})}{\sum_{i=0}^{n-1} X_i} + \pi.$$

We define $U_n = \sum_{i=0}^{n-1} X_i$, $\xi_i = Z_i - \pi X_{i-1}$ and $S_n = \sum_{i=1}^{n} \xi_i$. Note that U_n is $\mathcal{F}_{n-1}$-measurable, where $\mathcal{F}_{n-1}$ is the σ-algebra generated by $X_0, X_1, \ldots, X_{n-1}$, $Z_1, \ldots, Z_{n-1}$. Furthermore,

$$\begin{aligned} \sum_{i=1}^{\infty} U_i^{-2} \mathbb{E}(\xi_i^2 | \mathcal{F}_{i-1}) &= \sum_{i=1}^{\infty} \frac{\mathbb{E}((Z_i - \pi X_{i-1})^2 | \mathcal{F}_{i-1})}{(\sum_{j=0}^{i-1} X_j)^2} \\ &= \sum_{i=1}^{\infty} \frac{\pi(1-\pi) X_{i-1}}{(\sum_{j=0}^{i-1} X_j)^2} \leq \sum_{i=1}^{\infty} \frac{\pi(1-\pi)}{X_{i-1}}. \end{aligned}$$

This last sum is almost surely finite on A, because X_i is strictly positive and almost surely grows exponentially in i. So the set

$$\{\lim U_n = \infty, \sum_{i=1}^{\infty} U_i^{-2} \mathbb{E}(\xi_i^2 | \mathcal{F}_{i-1}) < \infty\}$$

contains A up to a set of measure zero. Now we may apply Lemma 3.4.1 and conclude that

$$\frac{\sum_{i=1}^{n} Z_i}{\sum_{i=0}^{n-1} X_i} \to 0 + \pi = \pi \qquad \text{a.s. on } A.$$

as $n \to \infty$.

Proof of Theorem 3.3.1(b). From Theorem 2.1 of [34] we have $(\sum_{i=1}^{n} X_i)/(\sum_{i=0}^{n-1} X_i) \to m$ a.s. on A. We apply Corollary 3.4.2, giving

$$\frac{\sum_{i=2}^{n+1} Z_i}{\sum_{i=1}^{n} Z_i} = \frac{(\sum_{i=2}^{n+1} Z_i)/(\sum_{i=1}^{n} X_i)}{(\sum_{i=1}^{n} Z_i)/(\sum_{i=0}^{n-1} X_i)} \frac{\sum_{i=1}^{n} X_i}{\sum_{i=0}^{n-1} X_i} \to \frac{\pi}{\pi} m = m, \qquad \text{a.s. on } A,$$

which proves the theorem.

The rate of convergence follows from Theorem 3.3.2. We use a part of Theorem 2.3 of [34] as a lemma to prove Theorem 3.3.2.

Lemma 3.4.3 *Assume that $m > 1$ and let Y be a standard normal random variable, independent of (X_n). For any x we have*

$$\mathbb{P}\Big[\frac{1}{\sigma}\Big(\sum_{i=1}^{n} X_{i-1}\Big)^{\frac{1}{2}}\Big(\frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_{i-1}} - m\Big) \le x | X_n > 0\Big] \to \mathbb{P}(Y \le x).$$

In the same way one can prove that

$$\mathbb{P}\Big[\frac{1}{\sqrt{\pi(1-\pi)}}\Big(\sum_{i=1}^{n} X_{i-1}\Big)^{\frac{1}{2}}\Big(\frac{\sum_{i=1}^{n} Z_i}{\sum_{i=1}^{n} X_{i-1}} - \pi\Big) \le x | X_n > 0\Big] \to \mathbb{P}(Y \le x).$$

Proof of Theorem 3.3.2. First we rewrite $\Big(\sum_{i=1}^{n} Z_i\Big)^{\frac{1}{2}}(\tilde{m}_n - m)$ as

$$\begin{aligned}
& \Big(\sum_{i=1}^{n} Z_i\Big)^{\frac{1}{2}}(\tilde{m}_n - m) \\
= \; & \Big(\sum_{i=1}^{n} Z_i\Big)^{-\frac{1}{2}} \sum_{i=1}^{n} (Z_{i+1} - mZ_i) \\
= \; & \Big(\sum_{i=1}^{n} Z_i\Big)^{-\frac{1}{2}} \Big[\sum_{i=1}^{n} (Z_{i+1} - \pi X_i) + \pi \sum_{i=1}^{n} (X_i - mX_{i-1}) \\
& \quad -m \sum_{i=1}^{n} (Z_i - \pi X_{i-1})\Big] \qquad (3.5)
\end{aligned}$$

We already know that on A, $(\sum_{i=1}^{n} Z_i)/(\sum_{i=1}^{n} X_{i-1})$ converges a.s. to the constant π and $(\sum_{i=1}^{n} X_i)/(\sum_{i=1}^{n} X_{i-1})$ converges a.s. to the constant m. Now the second term on the right-hand side of (3.5) can be rewritten as

$$\begin{aligned} & \pi\Big(\frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n X_{i-1}}\Big)^{-\frac{1}{2}}\Big(\sum_{i=1}^n X_{i-1}\Big)^{-\frac{1}{2}}\sum_{i=1}^n (X_i - mX_{i-1}) \\ = \quad & \pi\Big(\frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n X_{i-1}}\Big)^{-\frac{1}{2}}\Big(\sum_{i=1}^n X_{i-1}\Big)^{\frac{1}{2}}\Big(\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_{i-1}} - m\Big). \end{aligned}$$

From Lemma 3.4.3 we see that the second term converges in distribution to a normal distribution with zero mean and finite variance. We can treat the other terms on the right-hand side in the same way, which proves the theorem.

3.5 Estimating a second function of the parameters

Next we want to prove Theorem 3.3.1(c), which gives a consistent estimator for γ. We first do this for the special case where $m_- = \sigma_-^2 = 0$. After that we treat the general case, and finally we interpret the function of the parameters that we can estimate.

3.5.1 The case $m_- = \sigma_-^2 = 0$

If $m_- = \sigma_-^2 = 0$, then $\gamma = (1-\pi)m + \pi\sigma^2 + \pi m^2 + m^2$. Theorem 3.3.1(c) now reads:

Theorem 3.5.1 *As $n \to \infty$, we have*

$$n^{-1}\sum_{i=1}^n (Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - \tilde{m}_n\right)^2 \to (1-\pi)m + \pi\sigma^2 + \pi m^2 + m^2$$

in probability on A.

We first compute $\mathbb{E}\left[(Z_1+1)\left(\frac{Z_2}{Z_1+1} - m\right)^2 | X_0 = k\right]$. To do this, we remember (3.2):

$$\mathbb{E}\Big[\frac{1}{Z_1+1}|X_0 = k\Big] = \frac{1}{\pi(k+1)}(1-(1-\pi)^{k+1}); \tag{3.6}$$

and note that elementary computations yield

$$\mathbb{E}\Big[\frac{X_1}{Z_1+1}|X_0 = k\Big] = \frac{m}{\pi}(1-(1-\pi)^k); \tag{3.7}$$

$$\mathbb{E}\Big[\frac{(X_1)^2}{Z_1+1}|X_0 = k\Big] = \frac{m^2 k}{\pi}(1-(1-\pi)^{k-1}) + \frac{m^2+\sigma^2}{\pi}(1-(1-\pi)^k). \tag{3.8}$$

Now we can compute the desired expectation in a straightforward way:

$$\begin{aligned}
& \mathbb{E}\left[(Z_1+1)\Big(\frac{Z_2}{Z_1+1}-m\Big)^2|X_0=k\right] \\
= \;& \mathbb{E}\Big[\frac{\pi(1-\pi)X_1}{Z_1+1}|X_0=k\Big]+\mathbb{E}\Big[\frac{(\pi X_1-m(Z_1+1))^2}{Z_1+1}|X_0=k\Big] \\
= \;& ((1-\pi)m+\pi\sigma^2+\pi m^2)(1-(1-\pi)^k)+m^2(1-k\pi(1-\pi)^{k-1}) \\
= \;& ((1-\pi)m+\pi\sigma^2+\pi m^2)\mathbb{P}(Z_1\neq 0|X_0=k)+m^2\mathbb{P}(Z_1\neq 1|X_0=k).
\end{aligned}$$

With this expression in our hand, we can identify a suitable martingale. We denote by $\mathcal{F}_n$ the σ-algebra generated by $\{Z_i; 1\leq i\leq 2n\}$.

Lemma 3.5.2

$$\begin{aligned}
M_n \;:=\; & \sum_{j=1}^{n}\Big((Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 \\
& -\Big[((1-\pi)m+\pi\sigma^2+\pi m^2)\mathbb{1}_{\{Z_{2j-1}>0\}}+m^2\mathbb{1}_{\{Z_{2j-1}\neq 1\}}\Big]\Big)
\end{aligned}$$

is a martingale with respect to $\mathcal{F}_n$.

Proof. It is clear that M_n is measurable with respect to $\mathcal{F}_n$. Let $\xi_{n+1} := M_{n+1}-M_n$ be the increments and note that $\mathbb{E}(\xi_{n+1}|X_{2n},\mathcal{F}_n) = \mathbb{E}(\xi_{n+1}|X_{2n}) = 0$, where the last equality follows from the previous computation. Hence,

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = M_n+\mathbb{E}(\mathbb{E}(\xi_{n+1}|X_{2n},\mathcal{F}_n)|\mathcal{F}_n) = M_n.$$

Theorem 3.5.3 *We have, as* $n\to\infty$, *a.s. on* A,

$$n^{-1}\sum_{j=1}^{n}(Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 \to (1-\pi)m+\pi\sigma^2+\pi m^2+m^2.$$

Furthermore, writing

$$\tilde{S}_n(m) := \sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1}-m\right)^2,$$

we have a.s. on A,

$$n^{-1}\tilde{S}_n(m) \to (1-\pi)m + \pi\sigma^2 + \pi m^2 + m^2.$$

Proof. From Lemma 3.5.2 we have that M_n is a martingale with respect to $\mathcal{F}_n$, with increments,

$$\begin{aligned}\xi_j &:= (Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1} - m\right)^2 \\ &\quad - \left[((1-\pi)m + \pi\sigma^2 + \pi m^2)\mathbb{1}_{\{Z_{2j-1}>0\}} + m^2\mathbb{1}_{\{Z_{2j-1}\neq 1\}}\right].\end{aligned}$$

Now we apply Lemma 3.4.1 with the given ξ_j and $U_n = n$. On the set A, we have $U_n \to \infty$. To show that $\sum_{i=1}^{\infty} U_i^{-2}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) < \infty$ on A, we claim that there exists a constant $C < \infty$ such that

$$Var\left[(Z_1+1)\left(\frac{Z_2}{Z_1+1} - m\right)^2 |X_0 = k\right] < C,$$

uniformly in k. The computations that justify this claim are lengthy but straightforward, and are given in the appendix. Now we have

$$\begin{aligned}\mathbb{E}(\xi_i^2|\mathcal{F}_{i-1}) &= Var\left[(Z_{2i-1}+1)\left(\frac{Z_{2i}}{Z_{2i-1}+1} - m\right)^2 |\mathcal{F}_{i-1}\right] \\ &= \mathbb{E}\Big(Var\left[(Z_{2i-1}+1)\left(\frac{Z_{2i}}{Z_{2i-1}+1} - m\right)^2 |X_{2(i-1)}\right]|\mathcal{F}_{i-1}\Big) \\ &< C,\end{aligned}$$

and we conclude that

$$\frac{M_n}{n} \to 0.$$

Now write $n^{-1}\bar{M}_n = n^{-1}\sum_{j=1}^{n}\bar{\xi}_j$, where

$$\bar{\xi}_j = (Z_{2j}+1)\left(\frac{Z_{2j+1}}{Z_{2j}+1} - m\right)^2 - \left[((1-\pi)m + \pi\sigma^2 + \pi m^2)\mathbb{1}_{\{Z_{2j}>0\}} + m^2\mathbb{1}_{\{Z_{2j}\neq 1\}}\right]$$

are the martingale increments. Define $\bar{\mathcal{F}}_j$ as the σ-algebra generated by $\{Z_1, \ldots, Z_{2j+1}\}$. Now with the same arguments as for the a.s. convergence of

$n^{-1}M_n$ we may prove that $n^{-1}\bar{M}_n \to 0$ a.s. on A. Finally note that

$$\begin{aligned}\frac{1}{2n}(M_n + \bar{M}_n) &= \frac{1}{2n}\sum_{i=1}^{2n}(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1} - m\right)^2 \\ &\quad - \left[(1-\pi)m + \pi\sigma^2 + \pi m^2\right]\mathbb{1}_{\{Z_{2j-1}>0\}} - m^2\mathbb{1}_{\{Z_{2j-1}\neq 1\}},\end{aligned}$$

and the second result of the theorem follows.

We remark that

$$\sum_{i=1}^{n}(Z_i+1)(\frac{Z_{i+1}}{Z_i+1} - m)^2 - [(1-\pi)m + \pi\sigma^2 + \pi m^2\mathbb{1}_{\{Z_i>0\}} + m^2\mathbb{1}_{\{Z_i\neq 1\}}]$$

is not a martingale itself, so we can not use Lemma 3.4.1 directly.

Because we do not know m, we cannot use $\tilde{S}_n(m)$ for estimation purposes, and we also need to analyse the behaviour of $\tilde{S}_n(\tilde{m}_n)$. Some algebra yields

$$n^{-1}\left(\tilde{S}_n(m) - \tilde{S}_n(\tilde{m}_n)\right) = n^{-1}(m-\tilde{m}_n)^2\sum_{i=1}^{n} Z_i + (m^2 - \tilde{m}_n^2). \tag{3.9}$$

From Theorem 3.3.2 we know that the square root of $(m-\tilde{m}_n)^2\sum_{i=1}^n Z_i$ is the sum of three random variables, each converging in distribution to a normally distributed random variable with finite variance. So the square root of the first term on the right hand side converges in distribution to 0. Because 0 is a constant, the convergence is also in probability. If $A_n \to 0$ in probability, then $A_n^2 \to 0$ in probability, so $n^{-1}(m-\tilde{m}_n)^2\sum_{i=1}^n Z_i$ converges in probability to 0 on A. Together with Theorem 3.5.3 this proves Theorem 3.5.1.

3.5.2 The general case

Until now we considered the situation where the observed individuals have no further offspring. We now allow observed individuals to have some X-offspring in the generation after the observation. So in terms of epidemics, in this section we allow detected individuals to infect other individuals during the interval of detection. Theorem 3.3.1 gives us a consistent estimator (in probability) of $(1-\pi)m + \pi\sigma^2 + \pi m^2 + m^2 - 2\pi m m_-$ on the explosion set A.

To prove the theorem, one can compute that (we omit the lengthy details)

$$\begin{aligned}
&\mathbb{E}\Big[(Z_1+1)\Big(\frac{Z_2}{Z_1+1}-m\Big)^2 - \\
&[(1-\pi)m+\pi\sigma^2+\pi m^2]\mathbb{1}_{\{Z_1>0\}} - m^2\mathbb{1}_{\{Z_1\neq 1\}} + 2\pi m(m_-)\mathbb{1}_{\{Z_1>1\}}|X_0=k\Big] \\
&+\mathbb{E}\Big[\frac{\pi[\pi\sigma_-^2+(1-\pi)(m_-)]}{Z_1+1}\mathbb{1}_{\{Z_1>0\}} - \frac{\pi^2(m_-)^2}{Z_1+1}(\mathbb{1}_{\{Z_1>1\}}-\mathbb{1}_{\{Z_1=1\}})|X_0=k\Big] \\
&=0.
\end{aligned}$$

This leads to the following lemma, which can be proved as Lemma 3.5.2.

Lemma 3.5.4 *Let $\mathcal{F}_n$ be the σ-algebra generated by $\{Z_i; 1\le i\le 2n\}$. Then*

$$\begin{aligned}
M_n \;:=\; &\sum_{j=1}^{n}\Big[(Z_{2j-1}+1)\left(\frac{Z_{2j}}{Z_{2j-1}+1}-m\right)^2 - \\
&[(1-\pi)m+\pi\sigma^2+\pi m^2]\mathbb{1}_{\{Z_{2j-1}>0\}} - m^2\mathbb{1}_{\{Z_{2j-1}\neq 1\}} \\
&+2\pi m(m_-)\mathbb{1}_{\{Z_{2j-1}>1\}} + \frac{\pi[\pi\sigma_-^2+(1-\pi)(m_-)]}{Z_{2j-1}+1}\mathbb{1}_{\{Z_{2j-1}>0\}} \\
&-\frac{\pi^2(m_-)^2}{Z_{2j-1}+1}(\mathbb{1}_{\{Z_{2j-1}>1\}}-\mathbb{1}_{\{Z_{2j-1}=1\}})\Big]
\end{aligned}$$

is a martingale with respect to $\mathcal{F}_n$.

Using this lemma we now prove

Theorem 3.5.5 *Let γ and $\tilde{S}_n(m)$ be as in Theorem 3.3.1. As $n\to\infty$, we have a.s. on A,*

$$n^{-1}\sum_{j=1}^{n}(Z_{2j-1}+1)\Big(\frac{Z_{2j}}{Z_{2j-1}+1}-m\Big)^2\to\gamma.$$

Furthermore, we have a.s. on A,

$$n^{-1}\tilde{S}_n(m)\to\gamma.$$

We can prove this in the same way as we did for the special case of the previous subsection. The only extra thing to prove is that there exists a $C<\infty$ such that

$$Var\left[(Z_1+1)\left(\frac{Z_2}{Z_1+1}-m\right)^2|X_0=k\right]<C,$$

for all k in the general case too. Again the computations to justify this inequality are lengthy and straightforward, and are given in the appendix.

The argument to prove Theorem 3.3.1(c) can now be finished exactly as in the previous section.

It seems rather difficult to establish the rate of convergence for the estimator of the second function of parameters, but Theorem 3.3.3(a) gives a bound for this rate.

Proof of Theorem 3.3.3. From Corollary 3.1 in [35] it follows that if $\{S_{ni}, \mathcal{F}_{n,i}, 1 \leq i \leq n\}$ is a square integrable martingale array with differences $X_{ni} := S_{ni} - S_{ni-1}$ (where S_{n0} is defined as 0) and such that

$$\sum_{i=1}^{n} \mathbb{E}(X_{ni}^2 | \mathcal{F}_{n,i-1}) \to 0 \qquad \text{in probability,}$$

and $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$ for $1 \leq i \leq n$, then $S_{nn} = \sum_{i=1}^{n} X_{ni} \to 0$ in probability.

We use this with $S_{nn} = n^{-(\frac{1}{2}+\delta)} M_n$, where M_i is as in Lemma 3.5.4, and $X_{ni} = n^{-(\frac{1}{2}+\delta)} T_i$, where T_i are the summands of M_n. We define $\mathcal{F}_{n,i}, 1 \leq i \leq n$, as the σ-algebra generated by $\{Z_i; 1 \leq i \leq 2n\}$, so $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$ holds for $1 \leq i \leq n$. We have already shown that $\mathbb{E}(T_i^2) < C$ for some C, uniformly in i, hence

$$\sum_{i=1}^{n} \mathbb{E}(X_{ni}^2 | \mathcal{F}_{n,i-1}) \leq n^{-(1+2\delta)} nC \to 0 \qquad \text{in probability,}$$

so $n^{-(\frac{1}{2}+\delta)} M_n \to 0$ in probability. We can prove in the same way that $n^{-(\frac{1}{2}+\delta)} \bar{M}_n \to 0$, where $\bar{M}_n$ is as in the proof of Theorem 3.5.3. Now by the definitions of M_n and $\bar{M}_n$, we may see that on the explosion set A,

$$n^{\frac{1}{2}-\delta}(n^{-1}\tilde{S}_n(m) - \gamma) = n^{-(\frac{1}{2}+\delta)} M_n + n^{-(\frac{1}{2}+\delta)} \bar{M}_n + n^{-(\frac{1}{2}+\delta)} \sum_{i=1}^{n} f(Z_i),$$

where $f(x) = O(\frac{1}{x})$. Since $\frac{Z_{i+1}}{X_i} \to \pi$ a.s. on A and $m^{-i} X_i$ converges almost surely to a finite random variable, we know that $m^{-i} Z_i$ almost surely converges to an almost surely finite and positive random variable, $\bar{W}$ say. Now we use

the Toeplitz Lemma (Lemma 1.2 in [34]) to see that

$$\sum_{i=1}^{\infty} f(Z_i) < \left(\sum_{i=1}^{\infty} m^{-i}\right) C\frac{1}{\bar{W}} < \infty,$$

for some C. So $n^{\frac{1}{2}-\delta}(n^{-1}\tilde{S}_n(m) - \gamma) \to 0$ in probability on the explosion set A.

By using the rate of convergence of $\tilde{m}_n$, equation (3.9) and the arguments following that equation we see that for all $\delta_1 > 0$, $n^{-\delta_1}\tilde{S}_n(m) - n^{-\delta_1}\tilde{S}_n(\tilde{m}_n) \to 0$ in probability on A, and now with $\delta_1 = \frac{1}{2} + \delta$ the theorem follows.

3.5.3 Interpretation of γ

The expression

$$\gamma = (1-\pi)m + \pi\sigma^2 + \pi m^2 + m^2 - 2\pi m m_-$$

which appeared as the limit in the previous subsection, turns out to have a somewhat surprising interpretation: it appears as a second moment if we treat our Z-observations as a Galton-Watson process itself. To explain what we mean by this, we compute $\mathbb{P}(X_0 = l | Z_1 = k)$, when the a-priori distribution of X_0 is uniform on the integers between k and N, where $N >> k$. After this we let N tend to infinity.

$$\begin{aligned}\mathbb{P}(X_0 = l | Z_1 = k) &= \frac{\mathbb{P}(Z_1 = k | X_0 = l)\mathbb{P}(X_0 = l)}{\sum_{i=k}^{N}\mathbb{P}(Z_1 = k | X_0 = i)\mathbb{P}(X_0 = i)} \\ &= \frac{\mathbb{P}(Z_1 = k | X_0 = l)}{\sum_{i=k}^{N}\mathbb{P}(Z_1 = k | X_0 = i)}.\end{aligned}$$

First we compute the denominator for $N \to \infty$

$$\begin{aligned}\lim_{N\to\infty}\sum_{i=k}^{N}\mathbb{P}(Z_1 = k | X_0 = i) &= \sum_{i=k}^{\infty}\binom{i}{k}\pi^k(1-\pi)^{i-k} \\ &= \frac{1}{\pi}\sum_{j=0}^{\infty}\binom{k+j}{j}\pi^{k+1}(1-\pi)^j.\end{aligned}$$

The summands are exactly the probabilities of a negative binomial distribution, with parameters $k+1$ and π, so $\sum_{j=0}^{\infty}\binom{k+j}{j}\pi^{k+1}(1-\pi)^j = 1$. Therefore

we know that the denominator converges to π^{-1}.

With some abuse of notation, we write a superscript $*$ when we discuss probabilities and accompanying expectations after taking the limit for $N \to \infty$. This leads to

$$\mathbb{P}^*(X_0 = l | Z_1 = k) = \binom{l}{k} \pi^{k+1} (1-\pi)^{l-k}.$$

Now it is easy to compute that

$$\begin{aligned} \mathbb{E}^*(X_0 | Z_1 = k) &= \frac{k+1}{\pi} - 1, \\ Var^*(X_0 | Z_1 = k) &= \frac{(1-\pi)(k+1)}{\pi^2}. \end{aligned}$$

Now some straightforward computations yield

$$\begin{aligned} \mathbb{E}^*(Z_2 | Z_1 = k) &= (1-\pi)m_+ + mk, \\ Var^*(Z_2 | Z_1 = k) &= [(1-\pi)m + \pi\sigma^2 + (1+\pi)m^2 - 2\pi m m_-](k+1) \\ &\quad -(1-\pi)\pi m_- - \pi^2 \sigma_-^2. \end{aligned}$$

We see that $\mathbb{E}^*(Z_2 | Z_1 = k) = mk + O(1)$ for $k \to \infty$ and $Var^*(Z_2 | Z_1 = k) = \gamma k + O(1)$ for $k \to \infty$. In this sense, we again estimate a first and second moment, just as in the classical case where the full generation sizes are observed.

3.6 Estimating a third function of parameters

From Theorem 1.3 of [34] we know that we can estimate two moments of the offspring distribution of a Galton-Watson process and no other functions of the parameters consistently if only the generation sizes are observed. However, in our context Theorem 3.3.1 gives that we can under certain conditions estimate a third function of parameters.

We have already shown that for $k \to \infty$

$$\begin{aligned} \mathbb{E}\Big(\frac{Z_2}{Z_1 + 1} | X_0 = k \Big) &\to m, \\ \mathbb{E}\Big((Z_1 + 1) \left(\frac{Z_2}{Z_1 + 1} - m \right)^2 | X_0 = k \Big) &\to \gamma. \end{aligned}$$

We next compute $\mathbb{E}[(Z_1+1)(Z_3/(Z_1+1)-m^2)^2|X_0=k]$:

$$\begin{aligned}
& \mathbb{E}[(Z_1+1)\left(\tfrac{Z_3}{Z_1+1}-m^2\right)^2|X_0=k] \\
= \;& \mathbb{E}[(Z_1+1)\Big[\Big(\tfrac{Z_3-\pi X_2}{Z_1+1}\Big)^2+\pi^2\Big(\tfrac{X_2-mX_1}{Z_1+1}\Big)^2+m^2\Big(\tfrac{\pi X_1}{Z_1+1}-m\Big)^2\Big]|X_0=k] \\
= \;& \pi(1-\pi)m\mathbb{E}\Big(\tfrac{X_1}{Z_1+1}|X_0=k\Big)+\pi^2\sigma^2\mathbb{E}\Big(\tfrac{X_1}{Z_1+1}|X_0=k\Big) \\
& +m^2\mathbb{E}\Big((Z_1+1)\left(\tfrac{Z_2}{Z_1+1}-m\right)^2|X_0=k\Big)-m^2\pi(1-\pi)\mathbb{E}\Big(\tfrac{X_1}{Z_1+1}|X_0=k\Big) \\
= \;& ((1-\pi)m+\pi\sigma^2-m^2(1-\pi))\mathbb{E}\Big(\tfrac{Z_2}{Z_1+1}|X_0=k\Big) \\
& +m^2\mathbb{E}\Big((Z_1+1)\left(\tfrac{Z_2}{Z_1+1}-m\right)^2|X_0=k\Big).
\end{aligned}$$

Now note that on A,

$$\begin{aligned}
& \mathbb{E}[(Z_1+1)\Big(\frac{Z_3}{Z_1+1}-m^2\Big)^2|X_0=k] \\
= \;& ((1-\pi)m+\pi\sigma^2-m^2(1-\pi))\mathbb{E}\Big(\frac{Z_2}{Z_1+1}|X_0=k\Big) \\
& +m^2\mathbb{E}\Big((Z_1+1)\Big(\frac{Z_2}{Z_1+1}-m\Big)^2|X_0=k\Big) \\
\to \;& (\gamma-2m^2+2\pi m m_-)m+m^2\gamma \\
= \;& (m^2+m)\gamma-2(1-\pi)m^2m_+ \\
= \;& \gamma_*,
\end{aligned}$$

where the convergence is almost surely. Recall that

$$S_n^*(m):=\sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+2}}{Z_i+1}-m^2\right)^2.$$

We can use the same martingale argument as used for the proof of Theorem 3.5.5 to prove the following result.

Theorem 3.6.1 *The quantity*

$$n^{-1}\sum_{i=1}^{n}\Big[\Big(\Big[\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\Big]^2+\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\Big)(Z_i+1)\left(\frac{Z_{i+1}}{Z_i+1}-\frac{\sum_{j=2}^{n+1}Z_j}{\sum_{j=1}^{n}Z_j}\right)^2\Big]-n^{-1}S_n^*(m)$$

converges (on A) in probability to $2(1-\pi)m^2m_+$. Hence,

$$n^{-1}S_n^*(m)\to\gamma_*, \qquad \textit{in probability on } A.$$

In order to obtain an observable quantity (m is not), we need to bound $n^{-1}|S_n^*(\tilde{m}_n) - S_n^*(m)|$, we compute

$$
\begin{aligned}
& |\tilde{S}_n^*(m) - \tilde{S}_n^*(\tilde{m}_n)| \\
= \; & |\sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+2}}{Z_i+1} - m^2\right)^2 - \sum_{i=1}^{n}(Z_i+1)\left(\frac{Z_{i+2}}{Z_i+1} - (\tilde{m}_n)^2\right)^2| \\
= \; & |\sum_{i=1}^{n}\Big[(Z_i+1)(m^4 - (\tilde{m}_n)^4) - 2Z_{i+2}[m^2 - (\tilde{m}_n)^2]\Big]| \\
\leq \; & |\sum_{i=1}^{n}\Big[Z_i(m^4 - (\tilde{m}_n)^4) - 2\tilde{m}_{n+1}\tilde{m}_n Z_i[m^2 - (\tilde{m}_n)^2]\Big]| \\
& +|n(m^4 - (\tilde{m}_n)^4) + 2(Z_2 - \tilde{m}_{n+1}Z_1)[m^2 - (\tilde{m}_n)^2]| \\
\leq \; & 2|Z_2 - \tilde{m}_{n+1}Z_1||m^2 - (\tilde{m}_n)^2| + n|m^4 - (\tilde{m}_n)^4| \\
& +|(m^2 - (\tilde{m}_n)^2)^2\Big(\sum_{i=1}^{n} Z_i\Big) - 2(\tilde{m}_{n+1} - \tilde{m}_n)\tilde{m}_n[m^2 - (\tilde{m}_n)^2]\Big(\sum_{i=1}^{n} Z_i\Big)| \\
= \; & 2|Z_2 - \tilde{m}_{n+1}Z_1||m^2 - (\tilde{m}_n)^2| + n|m^4 - (\tilde{m}_n)^4| \\
& +|(m+\tilde{m}_n)(m-\tilde{m}_n)^3 + 2(m+\tilde{m}_n)\tilde{m}_n(m-\tilde{m}_{n+1})(m-\tilde{m}_n)|\Big(\sum_{i=1}^{n} Z_i\Big) \\
\leq \; & 2|Z_2 - \tilde{m}_{n+1}Z_1||m^2 - (\tilde{m}_n)^2| + n|m^4 - (\tilde{m}_n)^4| \\
& +|(m+\tilde{m}_n)(m-\tilde{m}_n)^3|\Big(\sum_{i=1}^{n} Z_i\Big) \\
& +|(m+\tilde{m}_n)\tilde{m}_n[(m-\tilde{m}_{n+1})^2 + (m-\tilde{m}_n)^2]|\Big(\sum_{i=1}^{n} Z_i\Big) \\
\leq \; & 2|Z_2 - \tilde{m}_{n+1}Z_1||m^2 - (\tilde{m}_n)^2| + n|m^4 - (\tilde{m}_n)^4| \\
& +|(m^2 - \tilde{m}_n^2)|(m-\tilde{m}_n)^2\Big(\sum_{i=1}^{n} Z_i\Big) \\
& +(m+\tilde{m}_n)\tilde{m}_n(m-\tilde{m}_n)^2\Big(\sum_{i=1}^{n} Z_i\Big) \\
& +(m+\tilde{m}_n)(m-\tilde{m}_{n+1})^2\Big(\sum_{i=1}^{n+1} Z_i\Big) - (m+\tilde{m}_n)(m-\tilde{m}_{n+1})^2 Z_1.
\end{aligned}
$$

Now $n^{-1}Z_1$ and $n^{-1}Z_2$ converge almost surely to 0. We can use the same arguments as used in Section 3.5 for the proof of convergence of $n^{-1}(S_n(\tilde{m}_n) - S_n(m))$ to see that $n^{-1}(S_n^*(\tilde{m}_n) - S_n^*(m))$ converges to zero in probability, which proves Theorem 3.3.1(d).

Note that $2(1-\pi)m^2m_+$ is not necessary in the plane spanned by m and γ, so we are able to estimate a third function of the parameters. Also note that if $m_- = 0$ we are unable to estimate the third parameter this way. However, if we look at the case where we observe a binomial distributed number of individuals from each generation, but observations do not influence the offspring distribution (that is, $m_+ = m_- = m$ and $\sigma_+^2 = \sigma_-^2 = \sigma^2$), we have an estimator for $(1-\pi)m^3$ and because we have an estimator for m that converges a.s., we are in theory able to estimate π consistently.

3.7 Conclusions

(A) From Theorem 1.3. of [34] we know that we cannot estimate more than two functions of the parameters (the first two moments) consistently only the generation sizes X_n of a branching process are given. In [38] it is showed that if we observe only a $Binomial(X_n, \pi)$ distributed fraction of the generation sizes, we can estimate two functions of parameters consistently, if π is known. In this chapter we have shown that, under certain conditions, we can estimate three functions of the parameters, even when we do not know π.

(B) For epidemiological purposes we want to estimate π as well, because this parameter gives an indication of how many individuals are infectious at a certain time, which may be important for implementing measures. In order to estimate this parameter in reasonable time we apparently need more and other information. We can possibly get this information by using contact tracing, i.e. finding out what contacts are made by an individual before it was observed and which contact may have caused the infection. Sometimes it is possible to get experimental information about the time between infection and removal of an individual, from that information we may also estimate π. Note that Becker and Hasofer [12] are able to estimate π and λ but they need information about the number of infectious individuals at the time of estimation, and this information is typically not available.

Appendix

In this appendix we show that for given i,

$$Var[(Z_1+1)(\frac{Z_{i+1}}{Z_1+1}-m^i)^2|X_0=k]$$

is bounded. This statement is equivalent to

$$\limsup_{k\to\infty}\mathbb{E}[(Z_1+1)^2(\frac{Z_{i+1}}{Z_1+1}-m^i)^4|X_0=k]<C$$

for all $i>0$ and some $C<\infty$.

$$\begin{aligned}
& \mathbb{E}[(Z_1+1)^2(\frac{Z_{i+1}}{Z_1+1}-m^i)^4|X_0=k] \\
= \;& \mathbb{E}[(Z_1+1)^{-2}[Z_{i+1}-m^i(Z_1+1)]^4|X_0=k] \\
= \;& \mathbb{E}[(Z_1+1)^{-2}[(Z_{i+1}-\pi m^{i-1}X_1)+m^{i-1}(\pi X_1-m[Z_1+1])]^4|X_0=k] \\
= \;& \mathbb{E}[(Z_1+1)^{-2}(Z_{i+1}-\pi m^{i-1}X_1)^4|X_0=k] \\
& +4\mathbb{E}[(Z_1+1)^{-2}(Z_{i+1}-\pi m^{i-1}X_1)^3m^{i-1}(\pi X_1-m[Z_1+1])|X_0=k] \\
& +6\mathbb{E}[(Z_1+1)^{-2}(Z_{i+1}-\pi m^{i-1}X_1)^2(m^{i-1}(\pi X_1-m[Z_1+1]))^2|X_0=k] \\
& +4\mathbb{E}[(Z_1+1)^{-2}(Z_{i+1}-\pi m^{i-1}X_1)(m^{i-1}(\pi X_1-m[Z_1+1]))^3|X_0=k] \\
& +\mathbb{E}[(Z_1+1)^{-2}(m^{i-1}(\pi X_1-m[Z_1+1]))^4|X_0=k].
\end{aligned}$$

We have assumed that $\mathbb{E}[(X_1-m)^4|X_0=1]<\infty$. Furthermore, we know that $\mathbb{E}[(Z_1-\pi)^4|X_0=1]<\infty$. We are dealing with branching processes. therefore, $\mathbb{E}[(Z_1+1)^{-2}(Z_{i+1}-\pi m^{i-1}X_1)^j|X_1=k]$ is the j-th central moment of the sum of k independent random variables with finite fourth moment and standard inequalities for sums of i.i.d. random variables give that there exists positive constants c_1, c_2 and c_3 such that

$$\begin{aligned}
\mathbb{E}[(Z_{i+1}-\pi m^{i-1}X_1)^4|X_1=k] &< c_1k^2, \\
\mathbb{E}[(Z_{i+1}-\pi m^{i-1}X_1)^3|X_1=k] &< c_2k, \\
\mathbb{E}[(Z_{i+1}-\pi m^{i-1}X_1)^2|X_1=k] &< c_3k. \qquad (3.10)
\end{aligned}$$

Taking these inequalities together brings us to

$$\begin{aligned}
& \mathbb{E}[(Z_1+1)^2(\frac{Z_{i+1}}{Z_1+1}-m^i)^4|X_0=k] \\
< \; & \mathbb{E}[(Z_1+1)^{-2}c_1(X_1)^2+4c_2X_1m^{i-1}(\pi X_1-m[Z_1+1])|X_0=k] \\
& +\mathbb{E}[(Z_1+1)^{-2}[6c_3X_1(m^{i-1}(\pi X_1-m[Z_1+1]))^2]|X_0=k] \\
& +\mathbb{E}[(Z_1+1)^{-2}(m^{i-1}(\pi X_1-m[Z_1+1]))^4|X_0=k].
\end{aligned}$$

Write

$$\begin{aligned}
\Delta &:= \pi X_1-m[Z_1+1] \\
&= \pi(X_1-m_+(X_0-Z_1)-m_-Z_1)+m_+(\pi X_0-Z_1)-m \\
&= \Delta_1+\Delta_2,
\end{aligned}$$

where $\Delta_1=\pi(X_1-m_+(X_0-Z_1)-m_-Z_1)$ and $\Delta_2=m_+(\pi X_0-Z_1)-m$. Observe that $\mathbb{E}(\Delta_1|X_0=k,\Delta_2=x)=0$ and $\mathbb{E}(\Delta_2|X_0=k)=-m$. Let Γ_+ be the third central moment of the offspring of an individual that is not detected and Γ_- of an individual that is detected, Similarly let ξ_+ be the fourth central moment of the offspring of an individual that is not detected and ξ_- of an individual that is detected. Basic computations yield:

$$\begin{aligned}
\mathbb{E}(\Delta|X_0=k) &= -m \\
\mathbb{E}\Big(\frac{\Delta^2}{Z_1+1}|X_0=k\Big) &= \mathbb{E}\Big(\frac{\Delta_1^2+\Delta_2^2}{Z_1+1}|X_0=k\Big) \\
&= \mathbb{E}\Big(\frac{\pi^2(X_0-Z_1)\sigma_+^2+\pi^2Z_1\sigma_-^2+\Delta_2^2}{Z_1+1}|X_0=k\Big) \\
\mathbb{E}\Big(\frac{\Delta^3}{(Z_1+1)^2}|X_0=k\Big) &= \mathbb{E}\Big(\frac{\Delta_1^3+3\Delta_1^2\Delta_2+\Delta_2^3}{(Z_1+1)^2}|X_0=k\Big) \\
&= \mathbb{E}\Big(\frac{\pi^3(X_0-Z_1)\Gamma_++\pi^3Z_1\Gamma_-}{(Z_1+1)^2}|X_0=k\Big) \\
&\quad +\mathbb{E}\Big(\frac{\pi^2[(X_0-Z_1)\sigma_+^2+Z_1\sigma_-^2]\Delta_2+\Delta_2^3}{(Z_1+1)^2}|X_0=k\Big)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\Big(\frac{\Delta^4}{(Z_1+1)^2}|X_0=k\Big) &= \mathbb{E}\Big(\frac{\Delta_1^4+4\Delta_1^3\Delta_2+6\Delta_1^2\Delta_2^2+\Delta_2^4}{(Z_1+1)^2}|X_0=k\Big)\\
&= \mathbb{E}\Big(\frac{\pi^4[(X_0-Z_1)\xi_+}{(Z_1+1)^2}|X_0=k\Big)\\
&\quad +\mathbb{E}\Big(\frac{3(X_0-Z_1)(X_0-Z_1-1)(\sigma_+^2)^2]}{(Z_1+1)^2}|X_0=k\Big)\\
&\quad +\mathbb{E}\Big(\frac{\pi^4[Z_1\xi_- +3Z_1(Z_1-1)(\sigma_-^2)^2]}{(Z_1+1)^2}|X_0=k\Big)\\
&\quad +\mathbb{E}\Big(\frac{[\pi^3(X_0-Z_1)\Gamma_+ +\pi^3Z_1\Gamma_-]\Delta_2}{(Z_1+1)^2}|X_0=k\Big)\\
&\quad +\mathbb{E}\Big(\frac{\pi^2[(X_0-Z_1)\sigma_+^2+Z_1\sigma_-^2]\Delta_2^2}{(Z_1+1)^2}|X_0=k\Big)\\
&\quad +\mathbb{E}\Big(\frac{\Delta_2^4}{(Z_1+1)^2}|X_0=k\Big).
\end{aligned}$$

To proceed we use that for all j:

$$\begin{aligned}
&\mathbb{E}\Big[\frac{\Delta_2^j}{(Z_1+1)^2}|X_0=k\Big]\\
\leq\ & 2\mathbb{E}\Big[\frac{\Delta_2^j}{(Z_1+1)(Z_1+2)}|X_0=k\Big]\\
=\ & 2\sum_{l=0}^{k}\binom{k}{l}\pi^l(1-\pi^{k-l})\frac{\Delta_2^j}{(l+1)(l+2)}\\
=\ & 2\sum_{l=0}^{k}\frac{1}{(k+1)(k+2)}\frac{1}{\pi^2}\binom{k+2}{l+2}\pi^{l+2}(1-\pi^{(k+2)-(l+2)})\Delta_2^j\\
\leq\ & 2\sum_{l=0}^{k+2}\frac{1}{(k+1)(k+2)}\frac{1}{\pi^2}\binom{k+2}{l}\pi^{l}(1-\pi^{k+2-l})\Delta_2^j\\
=\ & \frac{2}{(k+1)(k+2)}\mathbb{E}\Big[\Delta_2^j|X_0=k+2\Big].
\end{aligned}$$

Repeated use of this last inequality and inequalities like (3.10) implies that there exists a C such that

$$\limsup_{k\to\infty} \mathbb{E}[(Z_1+1)^2(\frac{Z_{i+1}}{Z_1+1} - m^i)^4 | X_0 = k] < C. \tag{3.11}$$

Chapter 4

On pair approximation

4.1 Introduction

Many models for describing the spread of an infection are based on the assumption of a randomly mixing population, where contacts between any two individuals are equally likely (see the previous chapters and e.g. [3, 27]). However, the assumption of random mixing is strong, therefore one may also want to consider epidemics on (social) networks. Connections in the network are possible contacts, e.g. if we consider sexually transmitted diseases and ignore all spread by other than sexual ways, the connections are only between people that have intercourse with each other.

We mainly consider an SIR description of infection spread, where individuals may be susceptible (S), infective (I) or removed/immune (R). A susceptible individual becomes infective if it has a contact with an infective individual and an infective individual becomes removed after some random time, which is distributed according to some given distribution. A removed individual never becomes susceptible or infective again. Contacts between two individuals that are connected to each other are made with rate τ. Births, deaths and migrations are ignored, i.e. we consider a closed population.

Our analysis is not restricted to this class of models. Other descriptions of infection spread are allowed. e.g. $SEIR$ dynamics, with latent/exposed class E, $SIRS$ dynamics and SIS dynamics, where recovered individuals become susceptible again. Note that if the infectious periods are exponentially distributed, the SIS process is the common contact process [53].

If the network, the initial states of the individuals and the relevant param-

eters of infectivity and infectious periods are known, in theory we can describe the epidemic process. However, analytic methods to describe the spread on general complex networks are not available. So it is necessary to use techniques that are not precise, but capture some features of the spread. One such a technique is using pair approximation (see e.g. [3, 10, 41, 43, 44, 62, 66, 67] for theory and [29, 30] for a computational intensive model on a real epidemic, using pair approximation techniques). The idea of pair approximations is basically the same as the idea behind randomly mixing populations, but here it are not the single individuals, but the pairs of connected individuals that are randomly mixing. In the next section we are more precise in defining the approximation. Out of the many different types of pair approximations, we follow Keeling [41] and Rand [67] in our discussion. The model presented in their papers is deterministic.

In this chapter we make the assumptions needed for pair approximation explicit. For ease of reference, we have numbered equations where approximation assumptions are made as $A1, A2, \ldots$. We start this chapter with some notation and basic assumptions. After that we explain the method of pair approximations. In Section 4.4 we discuss a new approximation of the reproduction number R_* (which is interpreted similarly as R_0). We also discuss other fundamental quantities like the rate of growth of the number of infective individuals, r and the critical infectivities τ_c and $\hat{\tau}_c$ (to be defined later). In Section 4.6 an infection on a given network is analysed rigorously, and results obtained by pair approximations are compared with exact results.

4.2 Set-up and notation

We consider a closed population of $\mathcal{N}$ individuals. We describe the population as a network $G = (V, E)$, where V is the set of individuals and E is the set of connections between individuals. We use $G(\mathcal{N})$ if we want to stress the dependence on the number of individuals in the network. If two individuals are connected, they are called neighbours of each other. Let n be the average number of neighbours of the individuals and n_v the number of neighbours of individual v. We assume that individuals can only infect their neighbours. In randomly mixing populations $n = \mathcal{N} - 1$.

As mentioned in the introduction of this chapter, we consider the SIR-epidemic. Let $[S] = [S](t)$ be the number of susceptible individuals at time t

and define $[I] = [I](t)$ and $[R] = [R](t)$ similarly. $[I](0) = I_0$, $[S](0) = \mathcal{N} - I_0$ and $[R](0) = 0$ are given. Because the population is closed $[S](t) + [I](t) + [R](t) = \mathcal{N}$. Let $[SI] = [SI](t)$ be the number of (ordered) pairs of neighbours of which the first is susceptible and the second is infective at time t. Note that under the random mixing assumption $[SI] = [S][I]$. More generally we define $[AB] = [AB](t)$ as the number of (ordered) pairs of neighbours of which the first is in state A and the second is in state B at time t, where A and B may be S, I or R.

A *triple* $v_iv_jv_k$ consists of three different individuals where v_j is connected to both v_i and v_k. A triple $v_iv_jv_k$ is a *triangle* if v_i is also connected to v_k. The position of an individual in the notation and the orientation of the triangle are important, so one physical triangle in the network is counted as six triangles in our notation and analysis. We write $[ABC] = [ABC](t)$ for the number of (ordered) triples of which the first individual is in state A, the second in state B and the third in state C at time t, where again A, B and C may be S, I or R. We distinguish between $[ABC]_\triangle$ and $[ABC]_\angle$, where the first denotes the number of triangles of individuals in states A, B and C, and the latter denotes the number of triples that do not form a triangle, with individuals in states A, B and C.

The parameter ϕ is defined as the number of triangles in a network divided by the number of triples. So

$$\phi := \frac{\sum [ABC]_\triangle}{\sum [ABC]}$$

where the sums are over all possible combinations of states A, B and C. This definition of ϕ cannot be used for infinite networks, therefore we introduce ϕ_v. Let ϕ_v be the fraction of pairs of neighbours of v that are neighbours of each other, i.e. ϕ_v is the number of triangles with v in the middle position, divided by the number of triples with v in the middle position. We note that on finite networks

$$\phi = \sum_{v \in V} \frac{n_v(n_v - 1)\phi_v}{\sum_{w \in V} n_w(n_w - 1)}. \tag{4.1}$$

We can extend the definition of ϕ to infinite regular graphs for which ϕ_v is the same for all v; for these regular networks we define $\phi := \phi_v$, which is consistent with (4.1).

If $[A](t), [B](t) > 0$, we define the "correlation" $C_{AB} = C_{AB}(t)$ as the fraction of the connected pairs that is in state AB, divided by the fraction of the individuals in state A times the fraction of the individuals in state B. In formula this reads

$$C_{AB} = \frac{[AB]/(n\mathcal{N})}{([A]/\mathcal{N})([B]/\mathcal{N})} = \frac{\mathcal{N}[AB]}{n[A][B]}. \tag{4.2}$$

For purposes becoming clear later, we also define

$$\bar{C}_{II} = \bar{C}_{II}(t) = [II](t)/(n[I](t)).$$

An infected individual has *infectious contacts* at rate $\tau = \beta/n$ with all of its neighbours independently of each other. If an infectious contact is made with a susceptible neighbour, the susceptible individual immediately becomes infective. If the neighbour is already infected or removed, nothing happens. If each individual has exactly n neighbours, then each individual makes contacts with rate β. Assume that the infectious period is exponentially distributed with parameter g. Choose $g = 1$ (which is always possible by rescaling time). Note that using constant rates implies the Markov property of the system.

We are interested in finding a quantity similar to the basic reproduction number R_0. Remember that R_0 is defined as the expected number of direct infections caused by one infective individual if all contacts are with susceptible individuals (see e.g. [27]). We do not use R_0 itself, because of the network structure: only for the first infected individual do we have that all possible first contacts are with a susceptible. However, we can define another quantity R_*, which is interpreted as the expected number of direct infections by one individual, infected after the start of the epidemic. In Section 4.4, we will give a more formal definition. It is still true that if $R_* \leq 1$, then the probability that a substantial part of the population will become infected (i.e. the probability that a major outbreak occurs) is 0; whereas $R_* > 1$ implies that this probability is positive.

We are also interested in the rate of growth of the number of infective individuals, r [27]. This r is only meaningful if the number of infective individuals grows or decreases exponentially, which holds in randomly mixing populations. Because we consider a network that is not complete, we use r_* for the rate of growth of the number of infective individuals on the network.

We will define it formally in Section 4.4.

The quantities R_* and r_* depend on τ. We define τ_c as $\inf\{\tau : r_* > 0\}$ and $\hat{\tau}_c$ as $\inf\{\tau : R_* > 1\}$. Those two critical infectivities are the same in randomly mixing populations and should be the same on networks, but because we need approximations to find R_* and r_*, the approximated τ_c and $\hat{\tau}_c$ may be different.

4.3 The approximations

4.3.1 Epidemics in randomly mixing populations

In this subsection we describe an epidemic in a randomly mixing population, so G is a complete network. We can write differential equations for the dynamics of $\mathbb{E}([S])$, $\mathbb{E}([I])$ and $\mathbb{E}([R])$:

$$\begin{array}{rclcl} \frac{d}{dt}\mathbb{E}([S]) & = & -\tau\mathbb{E}([SI]) & = & -\beta\frac{\mathbb{E}([SI])}{\mathcal{N}-1}, \\ \frac{d}{dt}\mathbb{E}([I]) & = & \tau\mathbb{E}([SI]) - \mathbb{E}([I]) & = & \beta\frac{\mathbb{E}([SI])}{\mathcal{N}-1} - \mathbb{E}([I]), \\ \frac{d}{dt}\mathbb{E}([R]) & = & & & \mathbb{E}([I]). \end{array} \tag{4.3}$$

This system of differential equations is not closed, because $\mathbb{E}([SI])$ is on the right-hand side but not on the left-hand side. A way to solve this problem is by approximating $\mathbb{E}([SI])$ by the expectation of singletons. Using

$$\mathbb{E}([SI]) = \mathbb{E}([S][I]) \approx \mathbb{E}([S])\mathbb{E}([I]) \tag{A1}$$

brings us back to one of the basic deterministic epidemic models [27]:

$$\begin{array}{rcl} \frac{d}{dt}\mathbb{E}([S]) & \approx & -\beta\frac{\mathbb{E}([S])\mathbb{E}([I])}{\mathcal{N}-1}, \\ \frac{d}{dt}\mathbb{E}([I]) & \approx & \beta\frac{\mathbb{E}([S])\mathbb{E}([I])}{\mathcal{N}-1} - \mathbb{E}([I]), \\ \frac{d}{dt}\mathbb{E}([R]) & = & \mathbb{E}([I]). \end{array} \tag{4.4}$$

If $\mathcal{N} \to \infty$, the dynamics of $\mathbb{E}([I])$ in a randomly mixing population is well described by the second differential equation (setting $\mathbb{E}([S])/(\mathcal{N}-1) = 1$).

This is consistent with the following argument: Consider the sequence of stochastic epidemic processes, $\{\mathcal{E}(\mathcal{N})\}_{\mathcal{N}\in\mathbb{N}}$, with a fixed number of initially

infected individuals. For any given time t, $[S](t)/(\mathcal{N}-1) \to 1$ almost surely, so the number of contacts made by infective individuals with other infective individuals before time t converges to 0 if $\mathcal{N} \to \infty$. This implies that at the start of the spread of an infection the number of infective individuals may with high probability be described by a birth and death process, with birth rate β and death rate 1. (For a more rigorous derivation of this result see [5]).

In a randomly mixing population with the contact structure described above, the basic reproduction number [27] is given by $R_0 = \beta$ and the rate of growth of the number of infective individuals is given by $r = \beta - 1$.

One remark; in many papers the quantity $\mathbb{E}([A])$ is expressed as $[A]$ and it is interpreted as the number of individuals in state A, where the notion of expectation is not even mentioned [27, 41]. Thinking of the dynamics of the epidemics, implicitly using expectations, but not making this explicit, may cause problems in interpreting the dynamics: expectations need not be integer, while the number of individuals in a certain state should be ([59] and the first paragraph of Section 1.3).

4.3.2 Pair approximation on networks: a first attempt

In this subsection, we consider the spread of an infection on a network that is not complete.

The differential equations describing the dynamics of $\mathbb{E}([S])$, $\mathbb{E}([I])$ and $\mathbb{E}([R])$, are given by (4.3). However, because of the network structure one might expect more correlation between the states of neighbours. Therefore the approximation (A1) is no longer considered adequate.

Instead of using an approximation for $\mathbb{E}([SI])$ in terms of singletons, it is also possible to take the time derivative of $\mathbb{E}([SI])$ and other pairs. We obtain [41, 67]

$$\begin{aligned} \frac{d}{dt}\mathbb{E}([SI]) &= \tau\mathbb{E}([SSI] - [ISI] - [SI]) - \mathbb{E}([SI]), \\ \frac{d}{dt}\mathbb{E}([II]) &= 2\tau\mathbb{E}([ISI] + [SI]) - 2\mathbb{E}([II]). \end{aligned} \tag{4.5}$$

We see that triples appear on the right-hand side. The differential equation describing the evolution of triples will have 4-tuples at the right-hand side and in general we need $(k+1)$-tuples to describe the evolution of k-tuples. The idea of pair approximation is that the number of triples in a certain state is

approximated in terms of the number of pairs and singletons in certain states.

A natural way to close the system at the level of pairs is as follows. Let $[A\cdot]$ be the number of pairs with an A-type individual in the first position and $[AB\cdot]$ be the number of triples with an A-type individual in the first and a B-type individual in the second position. The expected number of neighbours of an individual is n, therefore we approximate as follows;

$$[A\cdot] \approx n[A], \tag{A2}$$

$$[AB\cdot] \approx (n-1)[AB]. \tag{A3}$$

These approximations are exact if every individual has exactly n neighbours. In order to approximate $[ABC]$ we now assume that:

$$\frac{[ABC]}{[AB\cdot]} \approx \frac{[BC]}{[B\cdot]}, \tag{A4}$$

where the *number* of triples in a certain state is approximated, not only the expected number of triples in those states. In [41, 67] the even stronger approximation

$$[ABC] \approx \zeta \frac{[AB][BC]}{[B]} \tag{A5}$$

is used, where $\zeta = (n-1)/n$. The approximation is stronger, because $[AB]$ may be large.

The differential equations (4.5) can be approximated as

$$\begin{aligned} \frac{d}{dt}\mathbb{E}([SI]) &\approx \tau\mathbb{E}\Big(\zeta\frac{[SS][SI]-[IS]^2}{[S]} - [SI]\Big) - \mathbb{E}([SI]), \\ \frac{d}{dt}\mathbb{E}[II] &\approx 2\tau\mathbb{E}\Big(\zeta\frac{[IS]^2}{[S]} + [SI]\Big) - 2\mathbb{E}([II]). \end{aligned}$$

If we assume that the infection starts with no removed individuals and that the initial number of infective individuals is small compared to the population size, and if we regard the situation where $\mathcal{N} \to \infty$, then for any given time t, $[SS](t)/[S](t) \to n$ and $[SI](t)/[S](t) \to 0$ a.s. and in expectation, while $[SI](t) \nrightarrow \infty$ and $[I](t) \nrightarrow \infty$ a.s. and in expectation. Furthermore,

$$\Big(\frac{[SS]}{[S]} - n\Big)[SI] \to 0 \quad \text{a.s. and in expectation,} \tag{A6}$$

$$\frac{[IS]^2}{[S]} \to 0 \quad \text{a.s. and in expectation,} \tag{A7}$$

which implies

$$\frac{d}{dt}\mathbb{E}([SI]) \approx (n-2)\tau\mathbb{E}([SI]) - \mathbb{E}([SI]),$$

for large $\mathcal{N}$. Taking $\bar{\beta} = (n-2)\tau$ shows that under our assumptions the dynamics of $\mathbb{E}([SI])$ are described by the same equation (with a slightly changed parameter) as the dynamics of $\mathbb{E}([I])$ if random mixing is assumed.

4.3.3 Triangles in the network

Up to now we did not take the topology of the network into account. Loops in the network have been ignored, while these loops are present in many real networks: The probability that the friends of my friends are also my friends is larger than to be expected if links were established by pure chance alone. Also when individuals are not "mobile" (e.g. if the "individuals" are farms as in Chapter 2), it holds that if individual v_2 is close to both v_1 and v_3 then v_1 and v_3 are probably close too.

We need to approximate the number of triples in a certain state in terms of the number of individuals and the number of pairs. The arguments leading to (A4) and (A5) and the approximation that the number of triples that are not triangles in a certain state can be approximated by the fraction of triples that are not triangles times the number of triples in a certain state, i.e. $\mathbb{E}([ABC]_{\angle}) \approx (1-\phi)\mathbb{E}([ABC])$, bring us to

$$\mathbb{E}\Big(\frac{[ABC]_{\angle}}{[AB]}\Big) \approx (1-\phi)\zeta\mathbb{E}\Big(\frac{[BC]}{[B]}\Big) \tag{A8}$$

and

$$\mathbb{E}([ABC]_{\angle}) \approx (1-\phi)\zeta\mathbb{E}\Big(\frac{[AB][BC]}{[B]}\Big). \tag{A9}$$

Furthermore, we assume that the state of individual v_2 in $v_1v_2v_3$ does not depend strongly on the presence or absence of a connection between the individuals v_1 and v_3 [67]. This motivates the approximation

$$\frac{[ABC]_{\triangle}}{[A\cdot C]_{\triangle}} \approx \frac{[ABC]_{\angle}}{[A\cdot C]_{\angle}}. \tag{A10}$$

The individuals v_1 and v_3 in a triple that is not a triangle $v_1v_2v_3$ are not connected and therefore the dependence between their states is ignored. This leads to the approximation

$$[A\cdot C]_{\angle} \approx \frac{(1-\phi)n(n-1)\mathcal{N}[A][C]}{(\mathcal{N})^2} = \frac{(1-\phi)n(n-1)[A][C]}{\mathcal{N}}. \tag{A11}$$

Note that this assumption is not in general consistent with

$$[A\cdot C]_{\angle} = \sum_{B\in\{S,I,R\}} [ABC]_{\angle}.$$

The quantity $[A\cdot C]_{\triangle}$ is the number of triangles with an AC pair in it. A natural approximation is

$$[A\cdot C]_{\triangle} \approx [AC](n-1)\phi. \tag{A12}$$

If we take the stronger variant of (A10)

$$[ABC]_{\triangle} \approx \frac{[ABC]_{\angle}}{[A\cdot C]_{\angle}}[A\cdot C]_{\triangle}, \tag{A13}$$

this gives the approximation given in [41, 67]:

$$\mathbb{E}([ABC]_{\triangle}) \approx \phi\zeta\mathbb{E}\Big(\frac{\mathcal{N}[AB][BC][CA]}{n[A][B][C]}\Big) = \mathbb{E}\Big(\frac{[AB][BC]}{[B]}C_{AC}\Big). \tag{4.6}$$

If we use this, we obtain from the set of differential equations (4.3) and (4.5) the approximations

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}([SI]) &\approx \tau\mathbb{E}\Big[\zeta\frac{[SS][SI]}{[S]}((1-\phi)+\phi C_{SI}) - \zeta\frac{[IS]^2}{[S]}((1-\phi)+\phi C_{II}) - [SI]\Big] \\
&\quad -\mathbb{E}([SI]) \\
\frac{d}{dt}\mathbb{E}([II]) &\approx 2\tau\mathbb{E}\Big[\zeta\frac{[IS]^2}{[S]}((1-\phi)+\phi C_{II}) + [SI]\Big] - 2[II] \\
\frac{d}{dt}\mathbb{E}([I]) &= \tau\mathbb{E}([SI]) - \mathbb{E}([I]).
\end{aligned} \tag{4.7}$$

4.4 Computation of basic quantities

4.4.1 The reproduction number R_*

In this subsection we give a formal definition of the reproduction number R_*. However, in the setting of this chapter this definition cannot be applied immediately. Therefore, we also give an alternative definition $R_*^{(t)}$, which can be applied in the context of pair approximations. We assume that $R_*^{(t)} \approx R_*$ and therefore in the quest for R_* we compute and approximate $R_*^{(t)}$.

For the definition of R_* we consider a sequence of populations $\{G(\mathcal{N})\}$, with $\mathcal{N} \to \infty$ and structured in a given way (Constructing the sequence is straightforward in randomly mixing populations and in the random graphs constructed in the next chapter). The number of individuals infected after exactly k infection-steps in a population of size $\mathcal{N}$ is denoted by $I_k(\mathcal{N})$. We define

$$R_* := \limsup_{k\to\infty} \limsup_{\mathcal{N}\to\infty} (\mathbb{E}[I_k(\mathcal{N})])^{1/k}.$$

For infections spreading in randomly mixing populations or on random graphs the lim sup's can be replaced by lim's. The reproduction number can be interpreted as the asymptotic rate of growth of the number of infective individuals.

In the dynamical setting of this chapter, we need another definition. We obtain this alternative definition by using the following useful idea [10, 41]: We use differential equations similar to the equations given in the previous sections to describe the evolution of $\mathbb{E}(C_{SI}(t))$ and $\mathbb{E}(\bar{C}_{II}(t))$. We can find local minima (in time) for both of these expectations. Denote these minima by respectively x and y. We assume that the expectations $\mathbb{E}(C_{SI}(t))$ and $\mathbb{E}(\bar{C}_{II}(t))$ are only slowly departing from x and y. If $[I](t) = 0$ define $C_{SI}(t)$ as x and $\bar{C}_{II}(t)$ as y. We follow [41] and use the following approximations for computing $R_*^{(t)}$:

$$\begin{aligned} C_{SI} &\approx x, \\ \bar{C}_{II} &\approx y. \end{aligned} \tag{A14}$$

More formally, let A_t be the event that there are still infective individuals at time t and the first infection after time t happens before the first removal after time t. Let time $T = T(\epsilon)$ be the first time that $C_{SI} \leq x + \epsilon$, where $\epsilon > 0$ and usually chosen to be small. Now consider a sequence of networks with $\{G(\mathcal{N}) = G(\mathcal{N}, \phi_{\mathcal{N}}, n_{\mathcal{N}})\}_{\mathcal{N}\in N}$ where $N \in \mathbb{N}$ and $\mathcal{N} \to \infty$, $\phi_{\mathcal{N}} \to \phi$ and $n_{\mathcal{N}} \to n$. Let $X_{\mathcal{N}}(t)$ be defined as the total number of direct infections by the

first individual that is infected after time t on the graph $G(\mathcal{N})$. Define

$$\begin{aligned}\overline{R_*^{(t)}} &:= \limsup_{\mathcal{N}\to\infty} \mathbb{E}(X_{\mathcal{N}}(T)|A_T),\\ \underline{R_*^{(t)}} &:= \liminf_{\mathcal{N}\to\infty} \mathbb{E}(X_{\mathcal{N}}(T)|A_T).\end{aligned}$$

If the limit exists (i.e. if $\overline{R_*^{(t)}} = \underline{R_*^{(t)}}$), we define

$$R_*^{(t)} := \lim_{\mathcal{N}\to\infty} \mathbb{E}(X_{\mathcal{N}}(T)|A_T).$$

Note that R_* can be interpreted as the expected number of infections by an infected individual that is infected a long time after the introduction of the infection, while $R_*^{(t)}$ is defined as the expected number of infections by an individual, infected just after some stopping time.

We give an approximation of $R_*^{(t)}$ (and thus R_*), which seems quite natural. A newly infected individual (say v_2) at time t has approximately

$$\mathbb{E}\Big(\frac{[ISS]_{\angle}(t)}{[IS](t)}\Big)$$

neighbours to which its "infector" (say v_1) is not connected and

$$\mathbb{E}\Big(\frac{[ISS]_{\triangle}(t)}{[IS](t)}\Big)$$

neighbours to which v_1 is connected. Consider one of the susceptible neighbours (say v_3) of v_2. The probability that v_2 has at least one infectious contact with v_3 is $\tau/(\tau+1)$, because of the Markov property of the epidemic model. If v_3 is also connected to v_1, then the probability that there is a contact between v_1 and v_3 during the infectious period of v_1 but after the moment that v_2 is infected, is also $\tau/(\tau+1)$, by the Markov property. For the moment we ignore all other infectious contacts of infective individuals with v_3. If there are infectious contacts between v_3 and both v_1 and v_2, then we ascribe the infection to v_1. In fact the probability that the first of these two infectious contacts is with v_2, is $1/2$ but we ignore the probability that v_3 is infected by some other infective neighbour. In this way we hope to correct for that assumption. In Section 4.6 we show that on some regular networks this assumption is plausible. So we use the approximation

$$R_*^{(t)} \approx \mathbb{E}\Big(\frac{\tau}{\tau+1}\frac{[ISS]_\angle}{[IS]}\Big) + \mathbb{E}\Big(\frac{\tau}{\tau+1}\frac{1}{\tau+1}\frac{[ISS]_\triangle}{[IS]}\Big). \tag{4.8}$$

By the assumptions (A10) and (A12), we obtain

$$R_*^{(t)} \approx \frac{\tau}{\tau+1}\mathbb{E}\Big(\frac{[ISS]_\angle}{[IS]}\Big) + \frac{\tau}{(\tau+1)^2}\mathbb{E}\Big(\frac{(n-1)\phi[ISS]_\angle}{(1-\phi)n(n-1)(\mathcal{N})^{-1}[I][S]}\Big). \tag{4.9}$$

By using the same idea as used for approximation (A4) we write:

$$\begin{aligned} R_*^{(t)} &\approx \frac{\tau}{\tau+1}\mathbb{E}\Big((1-\phi)\zeta\frac{[SS]}{[S]}\Big) + \frac{\tau}{(\tau+1)^2}\mathbb{E}\Big(\frac{\zeta\phi[SS][IS]}{n(\mathcal{N})^{-1}[I][S]^2}\Big) \\ &= \frac{\tau}{\tau+1}\mathbb{E}\Big(\zeta\frac{[SS]}{[S]}\Big((1-\phi)+\frac{1}{\tau+1}\phi C_{SI}\Big)\Big). \end{aligned}$$

Recall that if $\mathcal{N}$ is large and the number of initial infective individuals is relatively small, then at the start of the epidemic $\mathcal{N}^{-1}[S] \approx 1$, $\mathcal{N}^{-1}[SS] \approx n$. Furthermore we have assumed in (A14) that $C_{SI} \approx x$. Using this, we obtain

$$R_*^{(t)} \approx \frac{\tau}{\tau+1}(n-1)(1-\phi) + \frac{\tau}{(\tau+1)^2}(n-1)\phi x. \tag{4.10}$$

The value of x may be deduced as follows. Compute the time derivatives of $\mathbb{E}(C_{SI})$ and $\mathbb{E}(\bar{C}_{II})$ under the assumption that $[I] > 1$. (Although Keeling does not use expectations, the computations are similar to [41]):

$$\begin{aligned} \frac{d}{dt}\mathbb{E}(C_{SI}) &= \tau\frac{\mathcal{N}}{n}\mathbb{E}\Big[\frac{[S][I]}{([S]-1)([I]-1)}\frac{[ISS]-[ISI]-[SI]}{[S][I]}\Big] \\ &\quad -\tau\frac{\mathcal{N}}{n}\mathbb{E}\Big[-\frac{[SI]^2([S]-[I]-1)}{[S]^2[I]^2}\Big], \\ \frac{d}{dt}\mathbb{E}(\bar{C}_{II}) &= \tau\mathbb{E}\Big[\frac{[I]}{[I]+1}\Big(2\frac{[ISI]+[IS]}{n[I]} - \frac{[IS][II]}{n[I]^2}\Big)\Big] - \mathbb{E}\Big[\frac{[I]}{[I]-1}\frac{[II]}{n[I]}\Big]. \end{aligned}$$

By using the assumptions (A3),(A4),(A10),(A11) and assuming that $\mathcal{N} \to \infty$ and $\mathcal{N}^{-1}[S] \to 1$ we obtain:

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}(C_{SI}) &= -\tau\mathbb{E}\Big[\frac{[I]}{[I]+1}\Big(C_{SI}+C_{SI}^2-(n-1)(C_{SI}-C_{SI}^2)(1-\phi)\Big)\Big] \\
&\quad -\tau\mathbb{E}\Big[\frac{[I]}{[I]+1}\Big(\phi(n-1)C_{SI}^2\bar{C}_{II}\Big)\Big], \\
\frac{d}{dt}\mathbb{E}(\bar{C}_{II}) &= \tau\mathbb{E}\Big[\frac{[I]}{[I]+1}\Big(2\phi(n-1)C_{SI}^2\bar{C}_{II}-nC_{SI}\bar{C}_{II}+2C_{SI}\Big)\Big] \\
&\quad -\mathbb{E}\Big[\frac{[I]}{[I]-1}\bar{C}_{II}\Big].
\end{aligned}$$

The quotients $[I]/([I]+1)$ and $[I]/([I]-1)$ are in the differential equations, because of the correlation between the changes in $[SI]$ and $[I]$. However, if the process does not go extinct $[I](t)$ will with high probability grow large, and therefore the quotients can be approximated by 1. Note that we assume (A14), so the right-hand sides are constants. We know that x and y are the fixed points of the equations, so

$$y = \frac{2\tau x}{1+n\tau x-2(n-1)\phi\tau x^2} \tag{4.11}$$

$$1 = (n-1)(1-\phi)(1-x)-\frac{2(n-1)\phi\tau x^2}{1+n\tau x-2(n-1)\phi\tau x^2}-x. \tag{4.12}$$

Therefore, the equation from which x can be deduced is:

$$n\zeta(1-x)(1-\phi)-\frac{2n\zeta\tau\phi x^2}{1+n\tau x-2n\zeta\tau\phi x^2}-x=1. \tag{4.13}$$

4.4.2 The rate of growth r_*

In a sequence of randomly mixing population $\{G(\mathcal{N})\}_{\mathcal{N}\in\mathbb{N}}$ the rate of growth r is defined as

$$r := \lim_{t\to\infty}\lim_{\mathcal{N}\to\infty}\frac{d}{dt}\log[\mathbb{E}([I](t))]. \tag{4.14}$$

The quantity can be of even more use than R_0 because it is relatively easy to estimate (Chapter 3). We also have the interpretation that if $r<0$, then a major outbreak will almost surely not occur, while if $r>0$, a major outbreak has positive probability.

In our context we do not assume random mixing, but in a large populations, we can still define some variant of r, by

$$r_* := \frac{d}{dt}\log[\mathbb{E}([I](T))], \tag{4.15}$$

where T is defined as in the previous section.

If $\mathcal{N}$ is large, the differential equations (4.7) can be further reduced to

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}([SI]) &\approx \tau\mathbb{E}\Big([SI]\Big[\zeta n(1-\phi)+\zeta n\phi C_{SI}-\zeta n\phi C_{SI}\bar{C}_{II}-1\Big]\Big)-\mathbb{E}([SI])\\
\frac{d}{dt}\mathbb{E}([II]) &\approx 2\tau\mathbb{E}\Big([SI]\Big[\zeta n\phi C_{SI}\bar{C}_{II}+1\Big]\Big)-2\mathbb{E}([II])\\
\frac{d}{dt}\mathbb{E}([I]) &= \tau\mathbb{E}([SI])-\mathbb{E}([I]).
\end{aligned} \tag{4.16}$$

For large $\mathcal{N}$, the last of these differential equations can be reformulated as

$$\frac{d}{dt}\mathbb{E}([I]) \approx n\tau\mathbb{E}([I]C_{SI})-\mathbb{E}([I]). \tag{4.17}$$

With $C_{SI}=x$ is constant (assumption (A14)), we obtain

$$r_* = \frac{d}{dt}\log(\mathbb{E}([I])) \approx n\tau x - 1. \tag{4.18}$$

The quantity r_* gives us an indication on whether or not the infection can spread in the population. So if $r_*>0$, spread is possible and if $r_*<0$, it has probability 0.

Remark: The estimates of R_* ($R_*^{(t)}$) and r_* are not consistent with each other. In randomly mixing populations $r>0 \Leftrightarrow R_0>1$, but this does not hold for R_* and r_*. However in the random mixing case and if $\phi=0$ the equivalence of $R_*>1$ and $r_*>0$ do hold.

4.4.3 The critical infectivities τ_c and $\hat{\tau}_c$

Let τ_c be the value of τ for which $r_*=0$, so τ_c is the value of τ for which $\tau n x=1$ holds. The definition of τ_c together with (4.13) gives:

$$\zeta(n-\frac{1}{\tau_c})(1-\phi)-\frac{2\zeta\phi\frac{1}{n\tau_c}}{1+1-2\zeta\phi\frac{1}{n\tau_c}}-\frac{1}{n\tau_c}=1. \tag{4.19}$$

We know by the remark at the end of the previous subsection that $r_* = 0$ does not necessarily imply that $R_* = 1$. We define $\hat{\tau}_c$ as the value of τ for which $R_* = 1$. So

$$\frac{\hat{\tau}_c}{\hat{\tau}_c + 1}(n-1)(1-\phi) + \frac{\hat{\tau}_c}{(\hat{\tau}_c + 1)^2}(n-1)\phi x = 1$$

We do not have an explicit expression for $\hat{\tau}_c$ based on this equality, but if we know n and ϕ. We can give a numerical value of $\hat{\tau}_c$.

4.5 Remarks

- The approximations (A9) and (4.6)

$$\begin{aligned} \mathbb{E}([ABC]_{\angle}) &\approx (1-\phi)\zeta\mathbb{E}\Big(\frac{[AB][BC]}{[B]}\Big) \\ \mathbb{E}([ABC]_{\triangle}) &\approx \phi\zeta\mathbb{E}\Big(\frac{\mathcal{N}[AB][BC][CA]}{n[A][B][C]}\Big), \end{aligned}$$

  may give nonsense: Consider $\mathbb{E}([III])$, which is approximated by

$$(1-\phi)\zeta\mathbb{E}\Big[\frac{[II]^2}{[I]}\Big] + \phi\zeta\mathbb{E}\Big[\frac{\mathcal{N}}{n}\Big(\frac{[II]}{[I]}\Big)^3\Big] = (1-\phi)\zeta n\mathbb{E}([II]\bar{C}_{II}) + \phi\zeta\mathcal{N}n^2\mathbb{E}(\bar{C}_{II}^3).$$

  For large $\mathcal{N}$, we have assumed that $\bar{C}_{II}$ stabilizes around y rather quickly. This y will be positive if there are still infective individuals with susceptible neighbours. Furthermore, y does not depend on $\mathcal{N}$ explicitly as long as $\mathcal{N}$ is large by (4.11). If $\phi > 0$ then the approximation for $[III]$ grows linearly in $\mathcal{N}$. So for large enough $\mathcal{N}$, the approximation for $[III]$ may be much larger than $n(n-1)[I]$, which is the real number of triples with an I in the centre position, which should be impossible.

  It is not clear what the consequences of this possibly failing approximation are for the analysis of SIR epidemics. The number of triples $[III]$ is not in (4.3) and the approximations for $[ISS]$ and $[IIS]$ (which are used in the differential equations) do not lead to such obvious problems, but it is not altogether clear that the approximations are good. However, the results of simulations in [41] are rather "consistent" with the predictions made using pair-approximations, so simulations suggest that the approximations are fit to describe SIR-dynamics.

Problems will arise if one uses pair-approximations for other more involved epidemic dynamics like in [8]. In the model described in that paper there are two states of infectivity. An individual can be mildly and severely infective. If a susceptible individual has a contact with an infective (mildly or severely) individual, it becomes mildly infective. If a mildly infective individual has a contact with another infective (mildly or severely) individual, it becomes severely infective. The contact rates of mildly and severely infective individuals may be different. Let $[I_m]$ be the number of mildly infective individuals and $[I_s]$ the number of severely infective individuals. The differential equation for the dynamics of $\mathbb{E}([I_s])$ has a term $\mathbb{E}([I_sI_m])$ on the right-hand side, while the differential equation for the dynamics of $\mathbb{E}([I_sI_m])$ has a term $\mathbb{E}([I_mI_mI_m])$ on the right-hand side. Pair approximation for $\mathbb{E}([I_mI_mI_m])$ will give inadequate results. This can be shown in the same way as the inadequacy of the approximation for $[III]$ in SIR-type epidemics is shown above.

- The approximations used for our computation are the expectations of (A4):

$$\mathbb{E}\Big(\frac{[ABC]}{[AB\cdot]}\Big) \approx \mathbb{E}\Big(\frac{[BC]}{[B\cdot]}\Big), \tag{4.20}$$

  the expectations of (A10):

$$\mathbb{E}\Big(\frac{[ABC]_\triangle}{[A\cdot C]_\triangle}\Big) \approx \mathbb{E}\Big(\frac{[ABC]_\angle}{[A\cdot C]_\angle}\Big),$$

  (A3), (A6), (A7), (A11), (A12) and (A14). Keeling and Rand [41, 67] are less precise in stating what their assumptions are, but all of the assumptions stated in this paper are implicitly used in their papers.

- The analysis of the spread of the infection by differential equations depends very strongly on the Markov-property of the spread. So if the infectious period has another distribution than the exponential distribution, or if the infectivity depends on the time since infection, the analysis changes and, as a rule, becomes much harder.

- It is generally not true that $\mathbb{E}([ABC]_\triangle) \approx \phi\mathbb{E}([ABC])$ throughout the epidemic. In the literature this equality seems to be used implicitly (e.g. [41], equation (10)), and it is the cornerstone for approximations as (A9) and (A11). The approximation is not used directly in (A10), which would make

analysis easier, but gives less reliable results. In [67] a correction is proposed, but this correction needs to be obtained from simulations, and differs from network to network.

- In [41] and in [10] the expression R_0 is used. This R_0 is not defined in the usual sense. Although the claim is that their R_0 means something similar as R_* in this chapter, it is in fact more like $r_* + 1$. In [41] $R_0 = xn\tau$. If $\phi = 0$ (which implies that $x = 1 - 2/n$ by (4.13)) and if τ is large, this R_0 may exceed n, which should be impossible, because only the neighbours of an individual can be infected by it. In [10] this problem is noticed and some ingenious methods are used to define an R_0 that cannot exceed n. However, Bauch is studying an SIS-type of infection. One can define R_0 as the expected number of times susceptible individuals become infective by a contact with the initial infective, during its first infective period. A-priori there is no reason that this does not exceed n, because neighbours may recover and become reinfected again. The way R_0 is computed in [10] also points in the direction of computing a quantity which can be directly obtained from r (this is mentioned by Bauch, but the interpretation of the computed R_0 stays vague). Note that, even with the definitions of R_0 proposed in [10] and [41], if C_{SI} really converges, the critical value $\hat{\tau}_c$ for which $R_* = 1$ is approximated correctly, because $n\tau[SI]/(n[I]) = 1$ implies $\tau[SI] = [I]$, which implies that the expected number of infectives after the next event (infection or removal) is the same as the number of infectives just before that event.
- An important observation is that C_{SI} stabilizes at a constant x by assumption and we may therefore write $\frac{d}{dt}[I] = \tau[SI] - [I] \approx n\tau x\frac{[S]}{\mathcal{N}}[I] - [I]$. This describes a mean-field model with infection rate $n\tau x$ and detection rate 1 (this is observed, but not explained in [44]).
- If $R_* > 1$, it is still possible that the infection goes extinct very quickly. The probability of a large outbreak cannot be obtained by the methods presented in this chapter, because all of the used approximations are about expectations.
- If the infection does not go extinct, the number of infected individuals is predicted to grow exponentially at the start of the epidemic. However, if the infection spreads on a regular two-dimensional lattice (e.g. the triangular lattice), it spreads like a travelling wave, so the number of infected

individuals will grow quadratically [58]. So, by using pair approximations one implicitly assumes that on the networks we use, there is exponential growth of the number of infectives at the start of the epidemic.

- It may be worth considering the invasory pair approximation idea of [10] in an SIR setting. Describing the dynamics will become harder than in the SIS-case, because of the extra state R.

4.6 An example

In this section we compare the R_* computed using pair approximation techniques, with the real R_* on a graph, where we can analytically obtain R_*. As before, we consider an SIR epidemic with exponentially (parameter 1) distributed infectious period and assume that an infective individual makes contact to every neighbour with rate τ. We choose an infinite network that is not finite dimensional in structure, because if the network is finite dimensional the number of infective individuals will grow polynomially [58]. Instead we choose a network $G = (V, E)$ that has a treelike structure and is transitive, i.e. for all $v_1 \in V$ and $v_2 \in V$ there exists an automorphism on G that maps v_1 on v_2. The network we take as an example can be seen as a tree of cliques (complete sub-networks). Every individual is member of two cliques of size four, where we assume that the cliques form a tree, i.e. the network has no loops apart from loops within a clique (See Figure 4.1).

We assume that there is one initial infective. In the network all individuals, apart from the initial infective, will cause the same expected number of direct infections, conditioned on the event that they are infected themselves. Now let the random variable X be the number of susceptible individuals in a clique containing the initial infective individual, that ultimately become infected. So X can take values $0, 1, 2, 3$. Let the random variable Y be the number of initially susceptible individuals in a clique containing the initial infective individual, that are directly infected by the initial individual. Note that it holds for every clique that if there has been an infective individual once, and there are no infective individuals anymore, then the infection will never be re-introduced in that clique because there are no loops in the network, apart from the loops within a clique. Furthermore, apart from the initial infective individual, every infected individual is the first infective in one clique and a secondary infective in another clique.

We may compute R_* for the infection on this network. We consider a newly infected individual, v_1. In the household of which v_1 is the first infective it will directly infect a random number of individuals, this random number is distributed as the random variable Y. The first infective individual in the household of which v_1 is a secondary individual is called v_0. The number of ultimately infected individuals in that particular household is X_{v_0} and the number of those infected individuals that are directly infected by v_0 is Y_{v_0}. Note that the pair (X_{v_0}, Y_{v_0}) is distributed as (X, Y). Because of "weighted" (or "size-biased") probabilities, the probability that the household in which v_1 is a secondary infective individual, totally has k secondary infective individuals is $(\mathbb{E}(X_{v_0}))^{-1}k\mathbb{P}(X_{v_0} = k)$. Because it is not possible to distinguish between the secondary individuals, the expected number of individuals directly infected by v_1, in the household of v_0 if the total number of secondary infections is $k \geq 1$, is given by $\mathbb{E}(k^{-1}(k - Y_{v_0})|X_{v_0} = k)$. So the expected number of individuals directly infected by v_1 is given by

$$\begin{aligned}
& \mathbb{E}(Y) + \sum_{k=1}^{4} \frac{k\mathbb{P}(X_{v_0} = k)}{\mathbb{E}(X_{v_0})} \mathbb{E}(k^{-1}(k - Y_{v_0})|X_{v_0} = k) \\
= \; & \mathbb{E}(Y) + \sum_{k=1}^{4} \sum_{l=0}^{k} \frac{\mathbb{P}(X = k)}{\mathbb{E}(X)} \mathbb{P}(Y = l|X = k)(k - l) \\
= \; & \mathbb{E}(Y) + \sum_{k=1}^{4} \sum_{l=0}^{k} \frac{1}{\mathbb{E}(X)} \mathbb{P}(Y = l, X = k)(k - l) \\
= \; & \mathbb{E}(Y) + \frac{1}{\mathbb{E}(X)} (\mathbb{E}(X) - \mathbb{E}(Y)) \\
= \; & \mathbb{E}(Y) + 1 - \frac{\mathbb{E}(Y)}{\mathbb{E}(X)}.
\end{aligned}$$

To compute the distribution of X the following theorem from [3] (originally in [4]) on epidemics in randomly mixing populations can be used:

Theorem 4.6.1 *Consider the standard SIR epidemic starting with $\bar{n}$ susceptible individuals and m infective individuals. Let λ be the total infection rate for an infective individual (i.e. $\lambda = \tau\bar{n}$) and $\mathcal{I}$ the infectious period. Denote by $P_k^{\bar{n}}$ the probability that the final size (without the initially infected individuals) of the epidemic is equal to k, $0 \leq k \leq \bar{n}$, and let $\phi(\theta) := \mathbb{E}[exp(-\theta\mathcal{I})]$ be the*

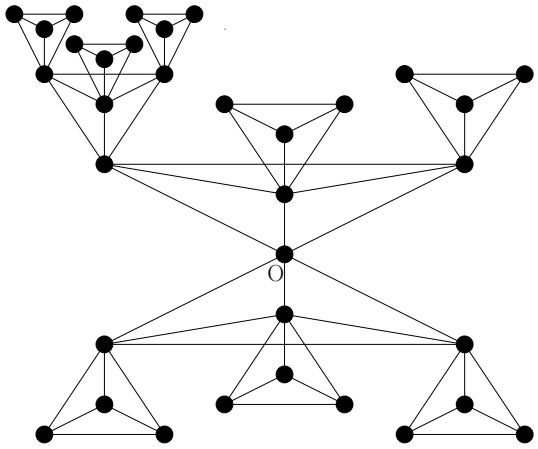


Figure 4.1: Network where each individual is member of two cliques of size four and where the cliques form a tree.

Laplace transform of $\mathcal{I}$. Then

$$\sum_{k=0}^{l} \binom{\bar{n}-k}{l-k} P_k^{\bar{n}} / [\phi(\lambda(\bar{n}-l)/\bar{n})]^{k+m} = \binom{\bar{n}}{l} \qquad \textit{for } 0 \le l \le \bar{n}.$$

If the infectious period is exponential with parameter 1, then

$$[\phi(\lambda(\bar{n}-l)/\bar{n})]^{k+m} = [\phi(\tau(\bar{n}-l))]^{k+m} = 1/[\tau(\bar{n}-l)+1]^{k+m}.$$

Note that we consider the spread in a clique of size 4, with $m = 1$ and $\bar{n} = 3$, so $\mathbb{P}(X = k) = P_k^3$. Now we can iteratively obtain $\mathbb{P}(X = k)$. By long but straightforward computations we obtain,

$$\mathbb{E}(Y) = \frac{\tau(6 + 54\tau + 180\tau^2 + 267\tau^3 + 175\tau^4 + 44\tau^5)}{2(1+\tau)^3(1+2\tau)^2(1+3\tau)}.$$

Using this we obtain for R_*:

$$\begin{aligned} R_* &= 1 + \frac{\lambda(6 + 54\lambda + 180\lambda^2 + 267\lambda^3 + 175\lambda^4 + 44\lambda^5)}{2(1+\lambda)^3(1+2\lambda)^2(1+3\lambda)} \\ &\quad - \frac{6 + 54\lambda + 180\lambda^2 + 267\lambda^3 + 175\lambda^4 + 44\lambda^5}{6(1 + 11\lambda + 45\lambda^2 + 73\lambda^3 + 48\lambda^4 + 12\lambda^5)}. \end{aligned} \tag{4.21}$$

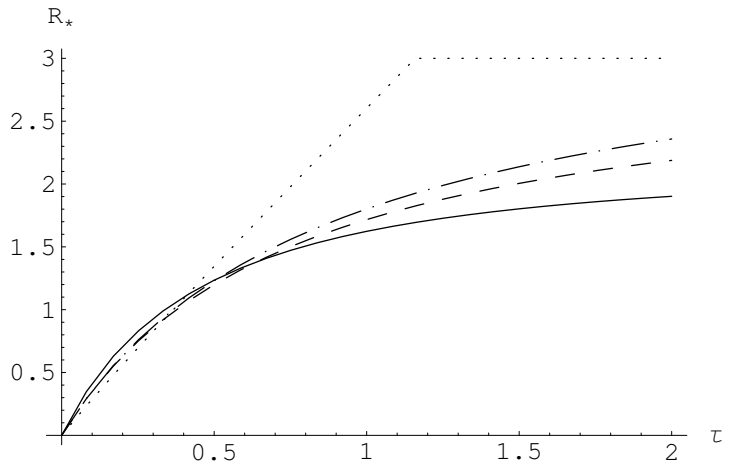


Figure 4.2: Plot of exact R_* (solid line), $R_*^{(t)}$ (dashed line), $\max(R_*^{(k)}, 3)$(dotted line) and $R_*^{(ka)}$ (alternating dot/dash line) against τ.

For the network in Figure 4.1, $\phi_v = 2/5$ for all v and $n = 6$. This is enough information to deduce x from (4.13).

We want to compare this exact R_* with the R_* obtained by pair approximations. In order to distinguish between the different approximations we use $R_*^{(k)}$ for the approximation of R_* proposed by Keeling [41], $R_*^{(ka)}$ for an alternative of this definition of Keeling, to be defined in the next paragraph, and $R_*^{(t)}$ for the new R_* approximation proposed in (4.10).

We already observed that $R_*^{(k)} = nx\tau$ explodes for large τ. An alternative definition of R_* could be $R_*^{(ka)} := nx\tau(1+\tau)^{-1}$, with the intuition that an infective individual has on average nx susceptible neighbours and the probability that there is an infectious contact with a given neighbour is $\tau(1+\tau)^{-1}$, but we also saw that under the assumptions of pair approximations $\tau_c = (nx)^{-1}$ is reasonable. In order to get $R_*^{(ka)} = 1$ if $\tau = \tau_c$, we change $R_*^{(k)}$ to

$$R_*^{(ka)} = \frac{nx\tau}{1+\tau}\frac{nx+1}{nx} = \frac{(nx+1)\tau}{1+\tau}. \tag{4.22}$$

The different R_* approximations are plotted for $0 \le \tau \le 2$ in Figure 4.2, where we have cut off $R_*^{(k)}$. In Figure 4.3 the differences between the approximated and the real R_* are given.

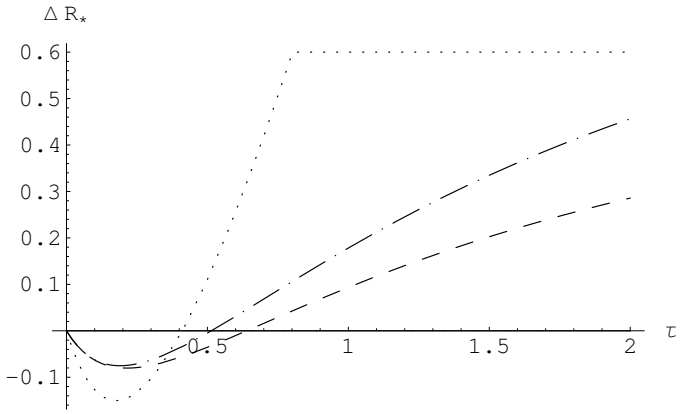


Figure 4.3: $R_*^{(t)} - R_*$ (dashed line), $\max(R_*^{(k)} - R_*, 0.6)$ (dotted line) and $R_*^{(ka)} - R_*$ (alternating dot/dash line).

We observe that especially for $\tau < 1$ the approximations of R_* are good. However, for large τ the approximations of R_* are too large. This is because we ascribe all secondary infected individuals in a clique, that has made infectious contacts with the first infective individual in that clique, to that first infective individual. In reality the spread of the infection can make detours, therefore we could have expected to obtain an estimated R_* that is too large.

From Figure 4.2 we see that the approximation for $\hat{\tau}_c$ (the value of τ for which $R_*^{(t)} = 1$) is rather good ($\hat{\tau}_c$, $\hat{\tau}_c^{(t)}$, $\hat{\tau}_c^{(k)}$ and $\hat{\tau}_c^{(ka)}$ are respectively approximated by 0.34, 0.38, 0.37 and 0.37, where the superscripts in $\hat{\tau}_c$ denote the approximation used.) We do not analyse r_* on this network, because it is hard to obtain the analytical results for this parameter on this network.

Suppose now that we look in retrospect to the spread of the infection on the network and our only observations are whether or not an infectious contact is made across a connection, but we do not know the times of the infectious contacts. We have to analyse the spread in another way: We say that a connection is open if an infectious contact is made across that connection, and otherwise it is closed. Define the open chemical distance of an individual to the initial infective individual as the minimal number of open connections that must be crossed to reach that particular individual from the initial infective and call all individuals at open chemical distance i from the initial infective individual generation i individuals. Let A_i be the event that there is a gener-

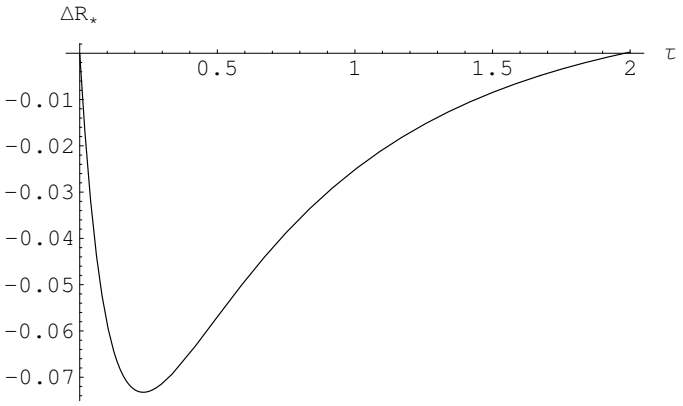


Figure 4.4: Difference of $R_*^{(t)}$, approximated as in (4.10), and $\hat{R}_*$, defined in (4.23)

ation i individual infected in the course of the epidemic. Let v_i be a uniformly chosen generation i individual that is infected in the course of the epidemic and let $\mathcal{R}(v_i)$ be the number of open connections of v_i to individuals that have no open connections to individuals at open chemical distance strictly less than i from the initial infective. Define $\hat{R}_* = \lim_{i\to\infty} \mathbb{E}(\mathcal{R}(v_i)|A_i)$.

Let Y be the number of individuals in a clique containing the initial infective individual that have an open connection to the initial infective. Because the marginal probability that a connection is open is $(\tau+1)^{-1}\tau$ and the number of individuals in a clique is 4, $\mathbb{E}(Y) = 3(\tau+1)^{-1}\tau$. By the same arguments as used for R_* we have

$$\hat{R}_* = \mathbb{E}(Y) + 1 - \frac{\mathbb{E}(Y)}{\mathbb{E}(X)} \tag{4.23}$$

We can compare this $\hat{R}_*$ with the approximation $R_*^{(t)}$. The difference of the approximated R_* (equation (4.10)) and $\hat{R}_*$ is given in Figure 4.4. We see that this difference is very small. We could have expected that $R_*^{(t)}$ is a better approximation for $\hat{R}_*$ than for R_*, because in the approximation (4.10) the way of ascribing infections to "infectors" is similar to the way this is done in the definition of $\hat{R}_*$.

Chapter 5

Random graphs as an approximation for networks on which infection spreads

5.1 Introduction

In Chapter 4 we presented one way to analyse the spread of an infection on a network, namely pair approximation. There are other ways to deal with epidemics on networks (see [43] for references).

If it is assumed that an infectious disease has special features such as fixed or exponentially distributed infectious periods and fixed infectivity, some features of the epidemic can be explored analytically on special networks such as regular lattices or trees (see [33, 39] for the needed theory.) However, for many networks, analytical methods to study the spread of the infection are not available.

Pair approximation as presented in the previous chapter has some drawbacks: In order to use pair-approximation techniques one needs to assume that the process has the lack-of-memory property, in the sense that all information needed to predict the future of the spread, are the number of susceptible, infective and removed individuals at the moment of prediction. In stochastic models, this assumption corresponds to exponentially distributed infectious periods of infected individuals and constant infectivity of the individual during that period. The lack-of-memory property is a rather strong assumption,

and therefore not desirable. Furthermore, the theory does not provide a way to estimate the probability of extinction of the infection.

Of course, computer simulations can be used to explore the dynamics of infection spread on networks, but a major disadvantage of simulations is that it is very hard to explain why the dynamics are as they are and on what properties of the graph and on what characteristics of the infection the overall dynamics depend. Furthermore, the parameter space may be too large to be sure that interesting behaviour of the spread does not occur in parts of the parameter space that are not explored.

As an alternative to pair-approximation, random graphs are proposed to describe the spread of the infection on large networks (see [28] for references). The network on which the infection spreads is replaced by a random graph with the same number of vertices as the original network and a degree distribution, D, that is based on the original network (e.g. one may choose $\mathbb{P}(D = k)$ to be the fraction of the individuals that have k neighbours in the original network). On this random graph we can take some randomness of the spread into account. In addition, we can compute or approximate the probability of extinction of the infection on the random graph.

However, only a small number of triangles arise naturally by the construction of the random graphs. Small loops in the network are important for the spread of the infection, because if two susceptible individuals, v_1 and v_2, that are neighbours of each other, have a common infected neighbour, v_3, they may both become infected by v_3. If that happens, v_1 cannot be infected by v_2 (and vice versa). So, the random graph models proposed in the literature [28], will overestimate the spread. The number of triangles in the network is used for pair approximations, but not in random graph models.

The main purpose of this chapter is to construct a random network to approximate a given network, that does have the desired expected number of triangles. The network consists of members of "households" (fully connected groups of individuals/cliques) and "bachelors" (single individuals, that are not part of any household). By choosing the ratio of households and bachelors, the distribution of the household sizes and the degree distribution of bachelors in a specific way, we can construct a random graph with the desired degree distribution and the desired fraction of "triples" that are also triangles.

One problem remains: It is very difficult to get analytic results for the spread of general infections on a random network. However, we can give

bounds for the reproduction number R_* (see Chapter 4) and the probability of extinction of the infection. Kuulasmaa [51] proved that it is possible to compare infections with the same marginal probability, say p, that at least one infectious contact between an infective and a given neighbour, is made during the infectious period (see also [25]). If the infectious periods are fixed and the infectivity is constant during this infectious period, R_* is maximal and the extinction probability of the infection is minimal. While if the infectious period is infinite with probability p, and 0 with probability $1-p$, R_* is minimal and the probability of extinction is maximal.

Analysing the spread of the "bounding infections" on the constructed random graph is not harder than analysing the spread of the infection by pair approximations, but now we can take the randomness of the spread into account. For the analysis of the epidemic on the random graphs we will use techniques from branching processes (see e.g. [39]).

In this chapter, an obvious relationship between the spread of an infection across a network and the theory of percolation is used (see Chapter 1 and [25]): Consider the original network and replace all connections between neighbours by two directed edges with opposite directions. Then ascribe positive random variables, $\mathcal{I}(v)$, to all individuals, where these random variables are i.i.d. and distributed as the infectious period of the infection, $\mathcal{I}$. Now the edge directed from v_1 to v_2 is closed with probability $e^{-\tau\iota(v_1)}$ (where $\iota(v_1)$ is a given realization of $\mathcal{I}(v_1)$) and open with probability $1-e^{-\tau\iota(v_1)}$, where τ is some constant.

This process can be coupled to a stochastic epidemic process. One can see this by interpreting the random variable $\mathcal{I}(v)$ as follows: If individual v becomes infective it stays infective for a random time $\mathcal{I}(v)$. The individuals that can be reached by paths from the initial infective individuals, are the individuals that ultimately become removed. If the infectious period is fixed at ι and $p := 1-e^{-\tau\iota}$ then the states of edges (open or closed) are independent of each other. This means that the random process described above is the same as the bond-percolation process on the directed network with parameter p [33]. If the infectious period is infinitely long with probability p and of duration 0 with probability $1-p$, the random process described above is linked to site percolation with parameter p: Vertices with infinite infectious period are open, while vertices with infectious period 0 are closed. The individuals that ultimately become removed are the individuals that are in the site-percolation clusters

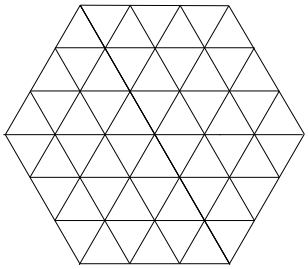


Figure 5.1: Triangular lattice

of the initial infective individuals, together with the neighbours of individuals in those clusters. On infinite connected graphs the critical probability p_c^{bond} is the infimum of the connection probabilities for which the probability that the origin is in an infinite bond percolation cluster is strictly positive. The critical probability p_c^{site} is defined similarly. (See Section 1.5).

In order to illustrate our analysis, we will consider the spread of an infection on the triangular lattice (Figure 5.1). For this lattice, p_c^{bond} and p_c^{site} are known explicitly. We may a-priori argue that our methods are not fit for analysing the spread of infections on such a regular network, but the triangular lattice can be used for demonstrating the method.

This chapter is organized as follows: In Section 5.2 some notation and terminology is introduced. In Section 5.3 we give results from percolation theory, which we use in the analysis of the epidemics. After that the construction of the random graphs is discussed, where the construction of the random graphs as previously done in the literature is our starting point. This construction leads to the construction of the random graphs with a "given expected fraction of connected triples that are also triangles". Finally we discuss the results and possible further research.

5.2 Notation and terminology

Some of the notation and terminology needed in this chapter is already given in Section 4.1. In this section we give further notation and terminology that we use for the definition of our networks and the description of the epidemic spread on the network.

First the words *network* and *graph* have the same meaning and we use the words depending on the context. In a same sense we speak of the *vertices* of a graph and the *individuals* in a network. *Edges* of the graph are *possible contacts* or *connections* between individuals of the network. Two individuals that are connected are *neighbours* of each other. We always assume that the number of vertices in the graph is large and that the average *degree* of a vertex (the number of edges with the given vertex as one of the endpoints) is small compared to the number of vertices. Later on we will be more precise by what is meant with "small" and "large". We consider graphs $G = (V, E)$, where V is the set of vertices and E is the set of edges in the graph. As in the previous chapter, $\mathcal{N} = |V|$ is the number of vertices in the graph. Let $\mathcal{N}_k$ be the number of vertices with degree k. Define n_v as the degree of individual v.

In the random graphs that are used in Section 5.4, the degree of v is a random variable $D(v)$, where the $D(v)$ are i.i.d. and distributed as the random variable D. The degrees of the individuals in the original network are considered to be samples from D, so D should be chosen in a proper way, e.g. the distribution may be defined as $\mathbb{P}(D = k) = \mathcal{N}^{-1}\mathcal{N}_k$. It is assumed that D has finite variance and that it does not depend on the number of vertices in the random graph. Vertices/individuals are denoted by v_i and edges/connections are denoted by v_iv_j, where v_i and v_j are the endpoints of the edge. We count v_iv_j and v_jv_i separately.

We again consider SIR-type systems in a closed population, i.e. systems where individuals may be susceptible (S), infected/infective (I) or removed/ immune/ death (R), and where, after removal, individuals never become susceptible or infective again. Birth and migration are ignored and deaths are only possible by removal. A dead individual is still part of the network, so the death of an individual does not influence the contacts between other individuals (this is contrary to assumptions in most mass-action models).

An infected individual has *infectious contacts* at rate τ with all of its neighbours. If an infectious contact is made with a susceptible neighbour, the susceptible individual immediately becomes infective. If the neighbour is already infected or removed, nothing happens. In this chapter we use p for the marginal probability that an infective individual v_1 has at least one infectious contact with a given neighbour v_2, so p is the marginal probability that an infected individual infects a given susceptible neighbour. We use $\mathcal{I}(v)$ for the

(random) length of the infectious period of individual v. The random variables $\mathcal{I}(v)$ are i.i.d. and distributed as a given random variable $\mathcal{I}$.

One remark: We do not need the assumption of a fixed infectivity during the infectious period, and of SIR-dynamics. We only use the probability that at least one infectious contact is made between an infectious individual and a neighbour. If we assume that there is a latent period ($SEIR$-dynamics) or that the infectivity depends on the time since infection, we can still compute the probability that individual v_1 has a contact with its neighbour v_2, during the infectious period of v_1.

In this chapter we use the formal definition of R_* from Chapter 4. So, we consider a sequence of populations with growing sizes and structured in a given way (Constructing the sequence is straightforward in randomly mixing populations and in the random graphs constructed in this chapter). The number of individuals infected at the k-th infection-step in a population of size $\mathcal{N}$ is denoted by $I_k(\mathcal{N})$. We define

$$R_* := \limsup_{k\to\infty} \limsup_{\mathcal{N}\to\infty} (\mathbb{E}[I_k(\mathcal{N})])^{1/k}.$$

For infections spreading in randomly mixing populations or on random graphs the lim sup's can be replaced by lim's.

On infinite graphs, we use $q = q(G, \mathcal{I}, p)$ for the probability of extinction of the infection, if the infection started with one uniformly chosen infected individual. The threshold probability $p_c = p_c(G, \mathcal{I})$ is defined by

$$p_c := \inf\{p : q(G, \mathcal{I}, p) < 1\}.$$

The theory and terminology of multi-type branching processes [39] is used in both the construction of random networks and in the spread of the infection. If there is only one initial infective individual, this individual will be the *root* or *ancestor*. Neighbours of this individual are in generation 1. Neighbours of generation i individuals that are not in generation j, for any $j \leq i$ are in generation $i + 1$. We use the maternal terminology, so we speak of mothers and daughters. In the networks we construct, an individual can have at most one mother. M is the mean offspring matrix. We use q_i for the probability that the branching process goes extinct if there is only one ancestor, which is of type i. Define $\bar{q} = (q_i, \cdots, q_r)$ as the probability of extinction vector of an

r-type branching process. Let s be a vector of length r. We always use $f^{(i)}(s)$ for the generating function of the number of daughters of a type i individual. The vector $f(s) = (f^{(1)}(s), \cdots f^{(r)}(s))$ is the vector of generating functions all with the same variable s.

In Section 5.4 we need a model, in which the ancestor has another offspring distribution than her progeny. Let $f(s)$ and q_i be defined as before where the offspring distribution is that of an individual that is the ancestor, let $\tilde{f}(s)$ be the vector of generating functions of the offspring distributions of individuals that are not the ancestor. Similarly, define $\tilde{q}_i$ as the probability that the offspring of an individual of type i, which is not the ancestor, goes extinct. Finally define $\tilde{q} = (\tilde{q}_i, \cdots, \tilde{q}_r)$.

5.3 Results from percolation theory

In the introduction of this chapter and in Chapter 1 we already discussed the relation between percolation theory and epidemics on networks. In Section 5.4 we use this relation to determine upper and lower bounds for R_*, q and p_c on random networks, if we do not know the actual distribution of the infectious period. We refer to [33] for an introduction to percolation theory.

We consider a percolation system that we will call locally dependent percolation [25, 51]: Assign a random variable $\mathcal{I}(v), 0 \leq \mathcal{I}(v) \leq \infty$ to each vertex v, where the random variables are i.i.d. and distributed as some random variable $\mathcal{I}$. Edges going out of v are open with probability $1 - e^{-\tau\iota(v)}$, where $\iota(v)$ is a realization of $\mathcal{I}(v)$ and τ is a positive constant. Conditioned on the realizations of $\mathcal{I}(v)$, the states of the edges are independent of each other. The marginal probability that an edge is open, is $p = \mathbb{E}(1 - e^{-\tau\mathcal{I}})$. The probability that a given edge v_1v_2 is open only depends on the realization of $\mathcal{I}(v_1)$: The state of v_1v_2 is correlated with the states of the other edges with v_1 as its first end vertex, but is independent of the states of the edges with another first vertex. The parameters τ, p and $\mathcal{I}$ can be interpreted in the way as they are defined in Section 5.2.

As noted in the introduction of this chapter, bond percolation and site percolation on directed graphs are special cases of locally dependent percolation. In [33] undirected graphs are considered. However, by [25] we know that the clusters of points that can be reached by open paths from a given vertex does not depend on whether the graph is directed or not. Particularly, the critical

(marginal) probability of connections

$$p_c = \inf\{p; \mathbb{P}(|\mathcal{C}_0| = \infty) > 0\}$$

does not depend on whether the graph is directed or not, where $\mathcal{C}_0$ is the set of vertices that can be reached by an open path for the origin. Note that in general p_c *does* depend on $\mathcal{I}$.

In [51] different epidemics of infections on a connected graph G are compared. The infections have different infectious periods, but the marginal probability p of having an infectious contact with a given neighbour is the same for all infections. Theorem 4.1 of [51] states implicitly that for given p, fixed infectious periods will give the smallest threshold value above which a major outbreak has positive probability. So the threshold for a major outbreak is above the critical value $p_c^{bond} = p_c^{bond}(G)$ of bond percolation on G. A direct consequence of Theorem 4.1(i) of [51] is that the infection with the fixed infectious period gives $R_* = R_*^{bond}$ which is larger than or equal to the R_* of any infection with the same marginal probability of at least one infectious contact across an edge. Furthermore, the infection with the fixed infectious period has a probability of extinction, q^{bond}, which is less than or equal to the probability of extinction of any other infection with the same marginal probability of infectious contacts.

Consider an infection, say A, such that with probability p an infectious individual will have infectious contacts with all of its neighbours and with probability $1-p$ it will have infectious contacts with none of his neighbours. A direct consequence of Theorem 2.1 of [51] (observed in [25]) is that A will have the largest threshold value above which a major outbreak has positive probability (i.e. for an infection with general infectious period this threshold will be less than the critical value $p_c^{site} = p_c^{site}(G)$ for site percolation on G.) Another direct consequence of [51] (but not mentioned before in the literature) is, that A also has the smallest R_* (R_*^{site}) and the largest probability of extinction (q^{site}) among all infections with marginal infection probability p.

So, for the critical probability p_c, the probability of extinction q and the reproduction number R_* holds that on every G

$$\begin{aligned} p_c^{bond} \le p_c(\mathcal{I}) \le p_c^{site}, \\ q^{bond} \le q(\mathcal{I}) \le q^{site}, \\ R_*^{site} \le R_*(\mathcal{I}) \le R_*^{bond}. \end{aligned} \tag{5.1}$$

Example: It is known [33] that on the triangular lattice the critical probabilities for bond and site percolation are given by

$$p_c^{bond} = 2\sin(\pi/18) \approx 0.347, \tag{5.2}$$
$$p_c^{site} = 1/2, \tag{5.3}$$

Hence, for the critical marginal infection probability p_c for infections spreading on the triangular lattice it holds that $2\sin(\pi/18) \leq p_c \leq 1/2$. Closed formulae for the extinction probabilities and R_* of epidemics on the triangular lattice are not known.

5.4 Random networks

5.4.1 Introduction to random networks

In the literature on pair approximations deterministic models to analyse the spread of an infection on a network are discussed. In this section we use a stochastic model. We approximate the given non-random network on which the epidemic spreads by a random network, which has a degree distribution "based" on the original network. The expected fraction of triangles, ϕ (see Chapter 4) in the random graph is the same as ϕ is the given non-random network. On the random network we have more tools to analyse the spread of the epidemic (see Chapter 2 of [28]). Throughout we assume that the number of individuals, $\mathcal{N}$ is large and that all individuals have a small (compared to $\mathcal{N}$) number of neighbours.

We use a sequence of random graphs to approximate the network. The limit behaviour for the epidemic as the number of vertices in the random graph grows to infinity, is analysed. The random graphs have a given expected ϕ and degree distribution, D, where $\mathbb{P}(D = k)$ can be chosen to be the fraction of the individuals in the original network, that have k neighbours. This gives insight in the spread of the epidemic on a graph with large $\mathcal{N}$. We use $\mathcal{N}^*$ for the number of vertices in the random graph.

5.4.2 Construction of an approximating network, not taking ϕ into account

The material in this subsection is entirely covered by [28], but we need the methods for the construction of random graphs with given ϕ. Here is how we construct a random graph $G(\mathcal{N}^*, D)$, with $\mathcal{N}^*$ vertices and degree distribution D:

- Let there be $\mathcal{N}^*$ vertices and assign a random number of so called *half-edges* (edges with only one endpoint assigned to a vertex) to each vertex, where the number of half-edges assigned to the vertices are i.i.d. and distributed as D.
- If the total number of half-edges is odd repeat the first step until the number of half-edges is even.
- Pair the half-edges at random (In such a way that all possible pairings have equal probability).

Because D has finite variance, the number of self loops (an edge that connects a vertex to itself) and the number of parallel edges (two edges with the same endpoints) in the constructed graph are asymptotically Poisson distributed with parameters independent of $\mathcal{N}^*$ (Theorem 2.1.1 of [28]). So if $\mathcal{N}^*$ is large, self loops and parallel edges are sparse in the network.

We consider a sequence of random graphs, $\{G(\mathcal{N}^*, D)\}_{\mathcal{N}^*}$, constructed as above, for a strictly increasing sequence, $\mathcal{N}^*$. For each of the graphs a vertex is uniformly chosen to be the origin (this individual at the origin will be the initially infected individual). The probability that the origin has degree k is $p_k := \mathbb{P}(D = k)$. The neighbours of the origin do not have degree distribution D, because it is k times as likely that a given vertex with k neighbours is a neighbour of the origin than that a given vertex with one neighbour is a neighbour of the origin (by the construction of the graph). Therefore the probability that an arbitrary neighbour of the initially infected individual has k neighbours is

$$\tilde{p}_k := \mathbb{P}(\tilde{D} = k) = (\mathbb{E}(D))^{-1} k \mathbb{P}(D = k), \tag{5.4}$$

where $\tilde{D}$ is a random variable with its distribution defined by the above equality. Other individuals that are infected during the epidemic also have the same degree distribution as $\tilde{D}$.

From Lemma A.2.2 of [36] we know that for large enough $\mathcal{N}^*$ and any $0 < \eta < 1/2$ the joint probability that there are loops in the first j generations (see Section 5.2) of the graph and the number of individuals in the first j generations does not exceed $(\mathcal{N}^*)^{\frac{1}{2}-\eta}$, is small. From this we can conclude that the start of an epidemic on the given random network will with high probability evolve the same as an epidemic on a random tree, where the tree itself is constructed by a branching process, with one individual in generation 0, which has k daughters with probability p_k. Other individuals than the ancestor have k daughters with probability $\tilde{p}_{k+1}$, where the "+1" is because one of the neighbours of such an individual is her mother. SIR-type systems on such a tree can be described by branching processes, for which many results are known [39].

5.4.3 Constructing the network with given expected ϕ

This is the key section of the chapter: The pair approximation techniques described in the epidemiological literature are mainly used in deterministic epidemic models, while the random graph methods of the previous subsection ignore small loops in the network. In this subsection we present a random-graph method of analysis of an epidemic, in which we also use the proportion of triples that is also a triangle, ϕ, of the network.

As before, if $\mathcal{N}$ is very large, the network G on which the epidemic spreads is replaced by a random network, on which we can analyse epidemic spread. We construct a sequence of random networks, $G(\mathcal{N}^*) = G(\mathcal{N}^*, D, \bar{D})$, where $\mathcal{N}^*$ is an increasing sequence and the random variable D and $\bar{D}$ are degree distributions, the meaning of $\mathcal{N}^*$ and the degree distributions is explained in the following paragraphs.

In the random network the expected fraction of triples that are triangles should be ϕ. To obtain this, we consider the vertices as super-individuals, which can be households (cliques, i.e. a fully connected group of individuals), or bachelors (single individuals, that are not part of any household). Each individual within a household has one neighbour outside its own household (so the number of individuals in a household is equal to the degree of the super-individual). We ignore all individuals without neighbours, because those individuals do not influence the epidemic. For convenience, households may be of size one, although in the graph they are exactly the same as bachelors

with one neighbour.

Let $\bar{\mathcal{N}}^*$, the number of super-individuals, be large and let $\bar{\mathcal{N}}_h^*$ be the number of households and $\bar{\mathcal{N}}_b^*$ the number of bachelors. Define $\gamma := (\bar{\mathcal{N}}^*)^{-1}\bar{\mathcal{N}}_h^*$, the fraction of households among the super individuals. The households are of random size, i.i.d. and distributed as $\bar{D}$. We define $\bar{p}_k := \mathbb{P}(\bar{D} = k)$. The outgoing edges (one from each individual in the household) are represented as half-edges. The degrees of the bachelors are i.i.d. and distributed as D, which is also the degree distribution of the original network G. Let $p_k = \mathbb{P}(D = k)$ as before. We assume that D has finite variance.

Because we want the degree distribution of the individuals (within or outside the households) also to be as the distribution of D, we need a relation between the distributions of D and $\bar{D}$. Each individual in a household of size k has k neighbours ($k-1$ within the household and 1 outside the household). Therefore it follows that

$$p_k = \frac{\gamma k\bar{p}_k + (1-\gamma)p_k}{\gamma\sum_{l=1}^{\infty} l\bar{p}_l + (1-\gamma)}. \tag{5.5}$$

So

$$p_k = \frac{k\bar{p}_k}{\sum_{l=1}^{\infty} l\bar{p}_l}. \tag{5.6}$$

We want to write $\sum_{l=1}^{\infty} l\bar{p}_l$ in terms of p_l. Observe that $\bar{p}_k = k^{-1}p_k\sum_{l=1}^{\infty} l\bar{p}_l$. By taking the sum over k on both sides, we get that $\sum_{l=1}^{\infty} l\bar{p}_l = (\mathbb{E}(D^{-1}))^{-1}$.

If individual v is a member of a household of size k, then $\phi_v = 1 - 2/k$. So one can compute ϕ as follows:

$$\begin{aligned}
\phi &= \sum_{l=2}^{\infty} \frac{\gamma l^2(l-1)\bar{p}_l(l-2)/l}{\gamma\sum_{k=2}^{\infty} k^2(k-1)\bar{p}_k + (1-\gamma)\sum_{k=2}^{\infty} k(k-1)p_k} \\
&= \frac{\gamma\sum_{k=2}^{\infty} k(k-1)(k-2)\bar{p}_k}{\gamma\sum_{k=2}^{\infty} k^2(k-1)\bar{p}_k + (1-\gamma)\sum_{k=2}^{\infty} k(k-1)p_k} \\
&= \frac{\gamma\sum_{k=2}^{\infty}(k-1)(k-2)p_k}{\sum_{k=2}^{\infty} k(k-1)p_k[\gamma + (1-\gamma)\mathbb{E}(D^{-1})]} \\
&= \Big(1 - 2\frac{\mathbb{E}(D)-1}{Var(D) + \mathbb{E}(D)(\mathbb{E}(D)-1)}\Big)\frac{\gamma}{\gamma + (1-\gamma)\mathbb{E}(D^{-1})}.
\end{aligned}$$

In this way we can construct random graphs with any $\bar{\mathcal{N}}^*$, any degree distribution with finite variance and any expected ϕ with

$$\phi \leq 1 - 2\frac{\mathbb{E}(D) - 1}{Var(D) + \mathbb{E}(D)(\mathbb{E}(D) - 1)}.$$

If the number of neighbours is fixed at n then

$$\phi = (n-2)\gamma/(\gamma n + 1 - \gamma). \tag{5.7}$$

Up to now we have only constructed the super-individuals of the network and not the network itself. The whole network can be constructed by connecting the super-individuals to each other in the same way as the individuals are connected in Subsection 5.4.2, where every household of size k has k half-edges assigned to it (1 from each individual in the household) and every bachelor of degree k has also k half-edges assigned to it.

Again we uniformly pick one individual as the origin, which can be an individual within a household or a bachelor. We construct a "tree with shortcuts" as in Section 5.4.2, the difference is that now we construct a tree of super-individuals. If the origin is a part of a household, this household is the root of the tree. If $\bar{\mathcal{N}}^* \to \infty$ the probability of shortcuts between the super-individuals in the first j generations of the tree goes to 0. So if $\bar{\mathcal{N}}^*$ is large enough, then with high probability, the start of an epidemic on the constructed network can be described as the start of an epidemic on a tree of super-individuals.

The randomly chosen root of this tree of super-individuals is a household of size k with probability

$$\frac{\gamma k \bar{p}_k}{(1-\gamma) + \gamma(\mathbb{E}(D^{-1}))^{-1}} = \frac{\gamma p_k}{(1-\gamma)\mathbb{E}(D^{-1}) + \gamma} \tag{5.8}$$

and a bachelor with k neighbours with probability

$$\frac{(1-\gamma)p_k}{(1-\gamma) + \gamma(\mathbb{E}(D^{-1}))^{-1}}. \tag{5.9}$$

The probability that a super-individual connected to the root is a household is given by:

$$\begin{aligned}\bar{\gamma} &:= \sum_k \frac{\gamma k \bar{p}_k}{\sum_l \gamma l \bar{p}_l + \sum_l (1-\gamma) l p_l} \\ &= \frac{\gamma (\mathbb{E}(D^{-1}))^{-1}}{\gamma (\mathbb{E}(D^{-1}))^{-1} + (1-\gamma)\mathbb{E}(D)} \\ &= \frac{\gamma}{\gamma + (1-\gamma)\mathbb{E}(D)\mathbb{E}(D^{-1})}.\end{aligned} \tag{5.10}$$

If a neighbour of the root is a household, it has size k, with probability

$$\frac{k\bar{p}_k}{\sum_l l\bar{p}_l} = p_k. \tag{5.11}$$

If a neighbour of the root is a bachelor it has k neighbours with probability

$$\frac{kp_k}{\sum_l lp_l} = \frac{kp_k}{\mathbb{E}(D)}. \tag{5.12}$$

This completes the construction of an approximating tree of super individuals. In Figure 5.2, a part of the tree is drawn. (Here the root is a household.)

5.4.4 Analysis of an epidemic on the approximating tree

The tree of super-individuals can also be seen as a tree of ordinary individuals with some shortcuts between "sisters". Throughout the rest of this section we use this paradigm and see the tree as a family-tree. We again speak of mothers, sisters and daughters. The infection is introduced in a household by only one individual (because we consider a tree there are no loops of super individuals), the other individuals in the household are all daughters of this one individual and are all connected to each other. We consider three types of ordinary individuals in the tree, bachelors, initial individuals of a household (the individual in a household that is connected to the lower generations in the tree of super-individuals, or the root if it is part of a household) and secondary individuals of a household (the other individuals in a household). These types of individuals are called respectively, type B, type H_1 and type H_2. In Figure 5.2 the individuals are marked by their type.

The root can only be of type B or type H_1. the daughters of type B individuals are of type B with probability $1-\bar{\gamma}$ and of type H_1 with probability $\bar{\gamma}$, independently of each other. The daughters of type H_1 individuals are

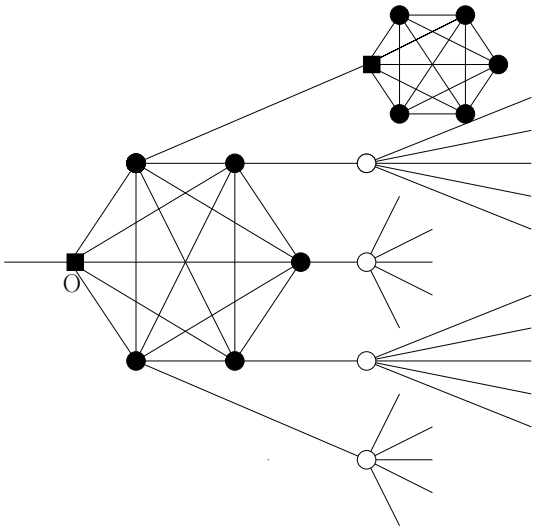


Figure 5.2: Random graph, for $p_6 = 1$ and $\phi = 2/5$, with the origin, B (open circles), H_1 (filled squares) and H_2 (filled circles) individuals marked.

always of type H_2, the only exception being the root. If the root is of type H_1, it will have one daughter of type H_1 (with probability $\bar{\gamma}$) or of type B (with probability $1-\bar{\gamma}$) all other daughters are of type H_2. H_2 individuals have exactly one daughter, which has type H_1 (with probability $\bar{\gamma}$) or type B (with probability $1-\bar{\gamma}$). If the root is a B-individual it has k daughters with probability p_k. All other B-individuals have $k-1$ daughters with probability $(\mathbb{E}(D))^{-1}kp_k$. An H_1-individual, that is the root has $k-1$ H_2-daughters with probability p_k and one daughter that is B (with probability $1-\bar{\gamma}$) or H_1 (with probability $\bar{\gamma}$). Other H_1 individuals have $k-1$ H_2-daughters with probability p_k.

We again consider two situations. First an infection with a fixed infectious period, second the infection for which the infectious period is infinite with probability p and 0 with probability $1-p$. So with the second type of infection all neighbours of an infective individual will become infective with probability p and none of the neighbours will be infected by the considered infective individual with probability $1-p$. The R_* for these infections are R_*^{bond} and R_*^{site} respectively (see Section 5.3).

It is very hard to compute R_*^{bond} in a correct way, because the dynamics of the spread within a household are hard to analyse. We change the model in the

following way: We ascribe the infection of all ultimately infected daughters to the mother. (Note that relations as "mother" and "daughter" are defined by the network not by the spread of the infection.) So if v_1 infects her daughter v_2 and a sister of v_2, v_3 is infected by v_2 then we say that v_3 is infected by v_1. It does not matter for the event that the infection goes extinct, whether we use the real infection paths or ascribe all infections to the mother. So neither p_c nor the probability of extinction changes by ascribing all infections to the mother. However the number of individuals that can be infected in n steps increases if all infections are ascribed to the mother (because in the real epidemic detours are made through households). This implies that if $R_* > 1$ the computed R_*^{bond} is larger than or equal to the real R_*^{bond}. This makes that R_*^{bond} is still an upper bound for R_* of infections with a marginal probability of infection p, but it may be less sharp.

Because of the reasons explained in the previous chapter, we ignore the root for computing R_*. We are only interested in the expected offspring size of individuals that are infected by another individual. A B individual infects an expected number of $p(1-\bar{\gamma})(\mathbb{E}(D^2)/\mathbb{E}(D)-1)$ other B individuals and $p\bar{\gamma}(\mathbb{E}(D^2)/\mathbb{E}(D)-1)$ H_1 individuals. The expected number of B individuals infected by a given H_2 individual is $p(1-\bar{\gamma})$ and the expected number of H_1 individuals infected by this individual is $p\bar{\gamma}$. There is no easy way to compute the expected number of H_2 individuals infected by one H_1 individual, but within a household all individuals are neighbours of each other, and the infection is introduced at only one individual within the household. So the epidemic within a household of size $\bar{n}$ can be viewed as an epidemic within a closed population starting with 1 infected individual and $\bar{n}-1$ susceptible individuals. To compute the distribution of the number of susceptible individuals that will ultimately become infected in this household Theorem 4.6.1 from the previous chapter is used.

If the infectious period is fixed at 1, then

$$[\phi(\lambda(\bar{n}-l)/\bar{n})]^{k+m} = exp[-(k+m)\lambda(\bar{n}-l)/\bar{n}] = (1-p)^{(k+m)(\bar{n}-l)}.$$

Now we can iteratively determine the distribution of the number of infective individuals ascribed to a H_1 individual. The mean of this distribution is $Z(p) := \sum_{\bar{n}} \sum_k p_{\bar{n}+1} k P_k^{\bar{n}}$. Therefore, the mean offspring matrix m^{bond} is known,

$$m^{bond} = \begin{pmatrix} p(1-\bar{\gamma})(\mathbb{E}(D^2)/\mathbb{E}(D)-1) & p\bar{\gamma}(\mathbb{E}(D^2)/\mathbb{E}(D)-1) & 0 \\ 0 & 0 & Z(p) \\ p(1-\bar{\gamma}) & p\bar{\gamma} & 0 \end{pmatrix}. \tag{5.13}$$

R_*^{bond} is the largest positive eigenvalue of m^{bond} [27, 39].

R_*^{site} is somewhat easier to calculate. Because if one of the daughters of a mother becomes infective, the infectious period of that mother was infinite, so all of its daughters are infected, with probability 1. This means that an H_1 individual infects no H_2 individuals with probability $1-p$ and all of its H_2 daughters with probability p. So the mean offspring matrix m^{site} is given by

$$m^{site} = \begin{pmatrix} p(1-\bar{\gamma})(\mathbb{E}(D^2)/\mathbb{E}(D)-1) & p\bar{\gamma}(\mathbb{E}(D^2)/\mathbb{E}(D)-1) & 0 \\ 0 & 0 & p(\mathbb{E}(D)-1) \\ p(1-\bar{\gamma}) & p\bar{\gamma} & 0 \end{pmatrix}, \tag{5.14}$$

which gives

$$R_*^{site} = \frac{p}{2}\Big[(1-\bar{\gamma})\Big(\frac{\mathbb{E}(D^2)}{\mathbb{E}(D)}-1\Big) + \sqrt{(1-\bar{\gamma})^2\Big(\frac{\mathbb{E}(D^2)}{\mathbb{E}(D)}-1\Big)^2 + 4\bar{\gamma}(\mathbb{E}(D)-1)}\Big]. \tag{5.15}$$

If $\mathbb{P}(D=n)=1$ then

$$R_*^{site} = \frac{p}{2}[(1-\gamma)(n-1) + \sqrt{(1-\gamma)^2(n-1)^2 + 4\gamma(n-1)}]. \tag{5.16}$$

In order to get the extinction probability for infections with fixed infectious periods we need to find $f_i^{bond}(s_1,s_2,s_3)$ for $1 \le i \le 3$ where the indices refer to respectively B, H_1 and H_2 individuals. After some algebra we get:

$$\begin{aligned} f_1^{bond}(s_1,s_2,s_3) &= \frac{\frac{d}{ds}f(s)}{\mathbb{E}(D)}\Big|_{s=1-p+s_1(1-\bar{\gamma})p+s_2\bar{\gamma}p}, \\ f_2^{bond}(s_1,s_2,s_3) &= \sum_{\bar{n}=0}^{\infty}\sum_{k=0}^{\bar{n}} p_{\bar{n}+1}P_k^{\bar{n}}s_3^k, \\ f_3^{bond}(s_1,s_2,s_3) &= s_1(1-\bar{\gamma})p + s_2\bar{\gamma}p + 1 - p, \end{aligned}$$

where P_k^n is as in Theorem 4.6.1 and $f(s)$ is the generating function of D.

The generating functions $f_i^{site}(s_1, s_2, s_3)$ for $1 \leq i \leq 3$ are given by

$$\begin{aligned} f_1^{site}(s_1, s_2, s_3) &= 1 - p + p\frac{\frac{d}{ds}f(s)}{\mathbb{E}(D)}|_{s=s_1(1-\bar{\gamma})+s_2\bar{\gamma}}, \\ f_2^{site}(s_1, s_2, s_3) &= 1 - p + p\frac{f(s_3)}{s_3}, \\ f_3^{site}(s_1, s_2, s_3) &= 1 - p + s_1(1-\bar{\gamma})p + s_2\bar{\gamma}p. \end{aligned}$$

Now the general introduction of this thesis and Chapter 4 of [39] can be used to find q_1, q_2 and q_3. Usually we can only find numerical values for these extinction probabilities.

5.4.5 An example

We consider the triangular lattice, so $p_6 = 1$ and $\phi = 2/5$. By (5.7) we have $\gamma = 1/5$, so (5.6) and (5.10) give $\bar{p}_6 = 1$ and $\bar{\gamma} = 1/5$. We can compute the expected offspring of an H_1 individual, $Z(p)$ exactly, but it is a complicated polynomial in p and we will not give it here. The mean offspring matrix (5.13) can be computed:

$$m^{bond} = \begin{pmatrix} 4p & p & 0 \\ 0 & 0 & Z(p) \\ 4p/5 & p/5 & 0 \end{pmatrix}.$$

The largest eigenvalue, R_*^{bond} of m^{bond} is given by

$$R_*^{bond} = 2p + \sqrt{4p^2 + \frac{pZ(p)}{5}} \leq 2p + \sqrt{4p^2 + p},$$

where the inequality is because $Z \leq 5$ by definition. Remember that p_c^{bond} is defined as the value of p for which $R_*^{bond} = 1$, which gives $p_c^{bond} \approx 0.225$. We use (5.15) to get

$$R_*^{site} = p(2 + \sqrt{5})$$

and $p_c^{site} = \frac{1}{2+\sqrt{5}} \approx 0.236$. The dependence of R_*^{bond} and R_*^{site} on p is given in Figure 5.3(a).

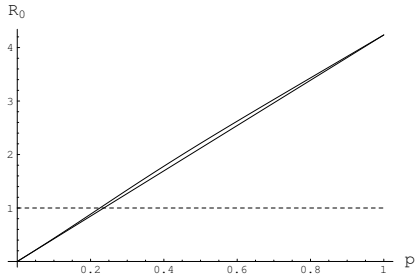
The generating functions are given by

$$\begin{aligned}
f_1^{bond}(s_1,s_2,s_3) &= (1-p+4/5s_1p+1/5s_2p)^5,\\
f_2^{bond}(s_1,s_2,s_3) &= \textstyle\sum_{k=0}^{5} P_k^5 s_3^k,\\
f_3^{bond}(s_1,s_2,s_3) &= 1-p+4/5s_1p+1/5s_2p\\
f_1^{site}(s_1,s_2,s_3) &= 1-p+p(s_1 4/5+s_2 1/5)^5,\\
f_2^{site}(s_1,s_2,s_3) &= 1-p+p(s_3)^5,\\
f_3^{site}(s_1,s_2,s_3) &= 1-p+s_1 4/5p+s_2 1/5p.
\end{aligned} \tag{5.17}$$

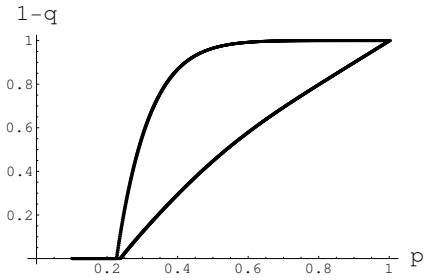
From this equation, we can numerically find the extinction probabilities. The probability that a uniformly chosen initial infective individual causes a major outbreak, $1-q$, is given in Figure 5.3(b).

5.4.6 Remarks

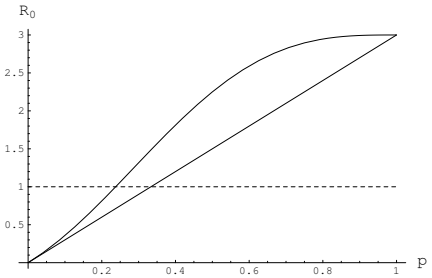
- A random graph is proposed as an approximating network of the actual network. If the infection does not go extinct, the number of infected individuals is predicted to grow exponentially at the start of the epidemic. However if the infection spreads on a regular two dimensional lattice (e.g. the triangular lattice) it spreads like a travelling wave, so the number of infected individuals will grow quadratically [58]. We did not succeed in finding a network (apart from the triangular lattice itself) with the right degree distribution and ϕ, on which the spread of an infection is also quadratic and for which it is possible to get analytic expressions for p_c. Note that pair approximations also predict exponentially growth.
- The number of vertices in the random graphs are random variables itself. It depends on the number of super individuals, $\bar{\mathcal{N}}^*$, the degree distributions of bachelors and the distribution of the household sizes.
- We can construct many networks with $\phi = 2/5$ and $n = 6$, that have exponentially growth at the start of an epidemic. Two examples of infinite networks for which $\phi_v = 2/5$ and $n_v = 6$ for all vertices v in the network are
  1. The network constructed in the previous chapter (Figure 4.1): The root has two groups of three daughters and the daughters from the same group are all connected to each other. Apart from the root all individuals have three daughters. All siblings are connected to each other.



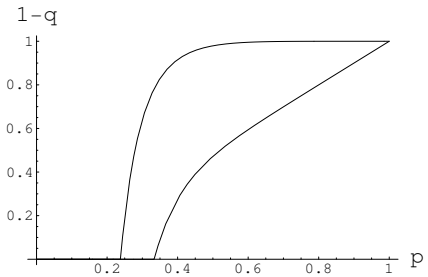
(a) R_* on random graph



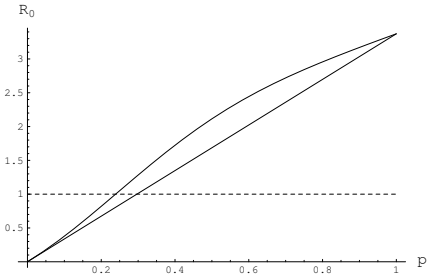
(b) survival probability of infection on random graph



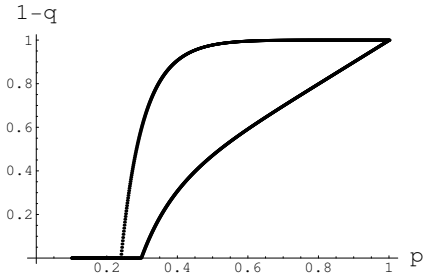
(c) R_* on graph where each individual is a part of two households of four individuals



(d) survival probability of infection on graph where each individual is a part of two households of four individuals



(e) R_* on graph where each individual is a part of one households of five individuals and has 2 neighbours outside the household



(f) survival probability on graph where each individual is a part of one households of five individuals and has 2 neighbours outside the household

Figure 5.3: The extreme reproduction ratios (R_*^{bond} and R_*^{site}) and the extreme survival probabilities ($1-q^{bond}$ and $1-q^{site}$) for different networks that approximate the triangular lattice, where the initial infective is chosen uniformly from all individuals in the random network.

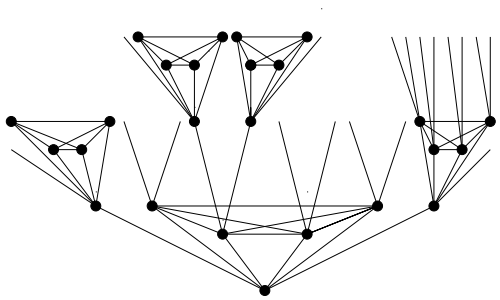


Figure 5.4: Tree where each individual is a member of a household of 5 individuals and has 2 neighbours outside the household.

2. The tree where the root has two single daughters and one group of four daughters. Within this group all daughters are connected to each other. Individuals that are not connected to sisters have one daughter not connected to sisters and four daughters in one group, where again all daughters within one group are connected to each other. All individuals that are a part of a group of siblings have two daughters, which are not connected to each other. (see Figure 5.4)

Because all individuals in these trees are topologically the same, we may choose any individual as the starting point of the epidemic.

In the first alternative network $R_*^{site} = 3p$ and therefore $p_c^{site} = 1/3$ and $R_*^{bond} = 3p(1 + 2p - 7p^3 + 7p^4 - 2p^5)$ by Theorem 3.6.1. This gives $p_c^{bond} \approx 0.238$. Because the two bounds are relatively far apart, in this network the distribution of the infectious period is important for determining $p_c(\mathcal{I})$ on this random tree.

In the second alternative network

$$m^{site} = \begin{pmatrix} p & 4p \\ 2p & 0 \end{pmatrix}, \tag{5.18}$$

which gives $R_*^{site} = (1+\sqrt{33})p/2$ and therefore $p_c^{site} = 2/(1+\sqrt{33}) \approx 0.297$. The matrix m^{bond} is harder to determine. In order to compute it, we first consider the spread of an infection with fixed infectious period on a network of five individuals, all connected to each other. Start with one infective individual and four susceptible ones. Let p and $\bar{Z}(p)$ be as before, but now

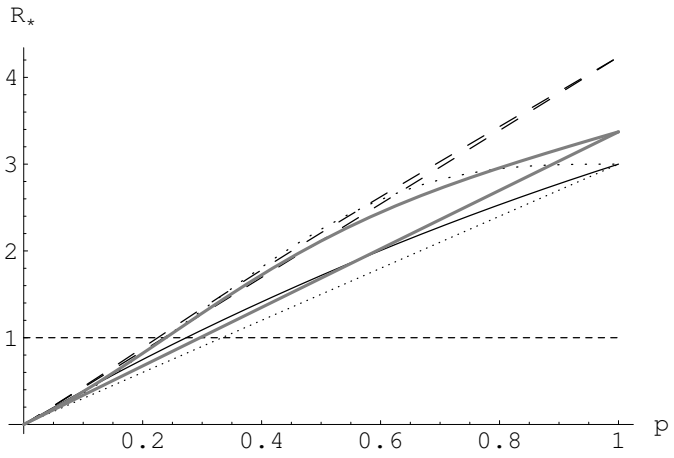


Figure 5.5: R_*^{bond} and R_*^{site} for the random tree (dashed lines), the tree where each individual is part of 2 households of size 4 (dotted lines), the tree where each individual is part of a household of size 5 and has 2 neighbours outside the household (solid grey lines) and R_* predicted by pair approximations as introduced in Chapter 4 (solid black line).

defined in households of size 5. By Theorem 4.6.1

$$m^{bond} = \begin{pmatrix} p & \bar{Z}(p) \\ 2p & 0 \end{pmatrix}, \tag{5.19}$$

which gives $R_*^{bond} = 1/2(p + \sqrt{p^2 + 8p\bar{Z}(p)})$, where $\bar{Z}(p)$ again is a complicated polynomial. The critical probability on this network is approximated by $p_c^{bond} \approx 0.239$. Again the difference between p_c^{bond} and p_c^{site} is considerable. So the actual distribution of the infectious period is important in order to get better estimates for the critical marginal probability $p_c(\mathcal{I})$. For the alternative networks, R_*^{bond} and R_*^{site} together with the probability of a major outbreak, if the infection started with one uniformly chosen infectious individual are given in Figure 5.3(c-f). To compare R_*^{bond} and R_*^{site} of different approximating networks and R_* predicted by pair approximations, we show them all together in Figure 5.5.

- It is possible to construct a network with $n = 6$ and $\phi = 2/5$ for which $p_c = 1$. Consider $G = (V, E)$ where $V = \mathbb{Z}$ and E are the edges connecting all vertices at distances 1, 3 or 4 of each other. Each individual has six

neighbours and $\phi_v = 2/5$ for all vertices v in the network. This network is one dimensional and if $p < 1$, with probability 1 there will be four vertices in a row at the right of the origin which have no infectious contacts to the right. And similarly with probability 1 there will be four vertices in a row at the left of the origin which have no infectious contacts to the left. Therefore a large outbreak is impossible on this network if $p < 1$.

- The message of the previous remark and earlier remarks is that we should have some a-priori idea of how the infection will spread before we can construct other networks on which we are capable of analysing the spread of infection. We assume that on most social networks there is exponential growth of the number of infectives at the start of the epidemic.
- On the triangular lattice the critical probability $p_c^* \approx 0.268$ found by pair approximation techniques is nearer to the p_c analytically obtained in percolation theory, than the critical values determined on random graphs or alternative graphs approximating the triangular lattice. Because of long range effects around the critical probability for percolation we could have expected that our estimates for p_c will not be very good: We have ignored all loops consisting of more than three vertices, by constructing a network that do not even have long loops. However, on networks that do have some tree-like structure (which is assumed to hold for social networks) the random graph method can be used.

5.5 Discussion

We have given a new way of analysing the spread of an infection on a network by using random graphs. Important factors like the distribution of the number of neighbours of individuals and the number of triangles in the network are used. The method presented seems to be as easy to use for computations as the differential equations obtained by pair-approximations. An advantage is that the random graphs allow us to deal with randomness in the spread of the infection, pair approximations only use the mean number of neighbours per individual. The full distribution of the number of neighbours can be used in the random graph method. The possibility to use a worst case scenario (the infection with a fixed infectious period) is a clear advantage of using random graph methods.

If we know the distribution of the infectious periods exactly it is still possible to use Theorem 4.6.1, but dependencies between the infectious period of individuals in the household and the number of individuals infected in that household make that it is not straightforward to compute R_* and the probability of extinction in this way. Ball pointed out [9] that it is possible to use [7] to compute R_* for general infectious periods.

One can use the model for the spread of epidemics between household proposed by Ball *et al.*[6], but this model is still a mass action model in which every household is connected to all other households. It is not clear how to construct a graph, with small degrees compared to the population size, that can be seen as an approximation of the network implicitly proposed for the household epidemics in [6].

In [64] other random graphs are considered, for which the expected ϕ (called clustering in [64]) and the mean degree can be computed as well. However, the construction of a network with a given expected ϕ and degree distribution is not made explicit there.

In reality it is usually not possible to observe the whole network on which an infection spreads, but the information from the network used in our network, ϕ and the degree distribution, can be estimated by observing only a part of the network.

It may be very hard to estimate parameters of the infection during the spread of an infectious disease itself (see Chapter 3), it may become even harder because we look at the epidemic from a "generation of infection" point of view, while real epidemics evolve in continuous time, where generations are mixing. Of course there is correlation between the number of infection steps needed to infect a given individual and the time at which the individual is infected, but it will be hard to incorporate this correlation in the estimates. It is worth further research how to estimate the parameters of the infection if we know the degree distribution and the ϕ of the network.

Another drawback of using random graphs and a "generation of infection" point of view is that it is not possible to use this way of analysis to describe the spread of infections in varying environments (Chapter 2), because in varying environments we need to keep track on the real time dynamics of the spread.

We did not study the final size of an infection, i.e. the number of individuals that will be removed at the end of the epidemic. It may be possible to obtain results for the final size [2], but it is questionable whether these results are

useful, because if $R_* > 1$ and a major outbreak occurs, then the final size has to be of the order of the population-size (this is proven for mass action interactions [3]). If such an infection spreads in the human population or in the population of farm animals, measures will be taken to stop the spread, which means that the environment is changing, some contacts will disappear and so on. All these changes are not incorporated in the model and the predicted final size will have no meaning in reality.

However, if the environment is not changing and if we consider the infection with fixed infectious period, then we can use the relation between bond percolation and the spread of the infection. Assume that there is only one open cluster of the same order as the population size. (This assumption holds for $p > p_c$ with high probability for the intersections of large boxes with $\mathbb{Z}^d$ [33], but not for all networks.) Because all contacts are symmetric the probability of individual v_1 not being affected by a major outbreak is the same as the probability that if v_1 is chosen as the initial infective a major outbreak will not occur. Because we chose the initial infective uniformly, the proportion of the population that is ultimately still susceptible will be the same as the probability of extinction of the infection.

The number of triangles in a network may be very important. On many regular lattices, small loops will occur for sure. The same holds for interaction networks of non-mobile individuals, where the probability of contacts is based on the distance between the two individuals. In some social networks the triangles seem to be present. In [1] and [73] the fraction of triangles are given of a network of film actors (where connections correspond to featuring in the same movie) and of networks of scientists (where connections correspond to being co-authors of at least one paper). These networks may not be the most probable networks on which infections spread, but they give some hints about the structure of other social networks.

Sexual and romantic networks are also investigated [11, 54]. These networks are important for the spread of sexually transmitted diseases. Because in a mainly heterosexual population the number of triangles is small, one would like to take the number of loops of length four or six into account. It turns out that even those small loops can be very rare, see for example the study done on romantic networks on an American high school [11], so there is some evidence that the spread of those disease may be described by random trees as presented in Subsection 5.4.2.

In this chapter we did not explicitly deal with power-law degree distributions, but at least some social networks have this property [1, 54, 73]. We say that the degree distribution has a power law if

$$p_k = c(k)k^{-\alpha} \qquad \text{for } k \in \mathbb{Z}_{>0} \text{ and } \alpha > 1, \tag{5.20}$$

where $\sum_{k\geq 0} p_k = 1$ and $\lim_{k\to\infty} c(k) = c > 0$. If we analyse the epidemic by approximating the network by the random networks constructed in Subsection 5.4.2, and if the degree distribution has a power law with $\alpha < 3$ then we predict $R_* = \infty$ if $\lambda \neq 0$. This is because the expected degree of an individual in the first generation is infinite.

If we construct a random network as in Subsection 5.4.3, we may get totally different results. Let the degree distribution have a power law with $2 < \alpha < 3$ and if all super-individuals are households ($\gamma = 1$), then the degree of the super-individuals is distributed with a power law with parameter $\alpha + 1$. The expected number of neighbours of a super-individual in generation 1 is finite, so if the infectivity is positive but small enough, R_* will be less than 1. So if we only use the degree distribution, we predict that a major outbreak has positive probability, while if we use the network of households the probability that many households are infected may be 0. In this example $\phi = 1$. It is an interesting open question whether it is possible to construct a network with a degree distribution following a power law with $2 < \alpha < 3$ and $\phi < 1$ such that the infection has non-zero infectivity, but $R_* < 1$.

Chapter 6

Infection spread in a population of individuals with random infectivity and susceptibility

6.1 Introduction

Most of the models for the spread of an infection are based on the assumption that all individuals have the same susceptibility and infectivity and, moreover, those models do not differentiate between susceptibility and infectivity and combine them in a single transmission rate parameter. This means that if a susceptible individual has a contact with an infectious individual it has a given probability that it becomes infective itself, where this probability is the same for every individual and for every contact with an infectious individual. On the other hand if an individual is infected, it contacts other individuals with a fixed rate, which is the same for every individual (See e.g. the other chapters of this thesis and [3, 27]). This assumption can be relaxed by allowing for several types of individuals (e.g. individuals may be differing in age, sex, species), where different types of individuals may have different infectivities and susceptibilities (these models are also discussed in [3, 27]), but still all individuals of a given type have the same infectivity and susceptibility.

In "reality" infectivity and susceptibility will show individual variation beyond the variation of characteristics, notably because of immunological polymorphism, or due to polymorphic reactions to vaccination [13, 14]. Here we explore a way to take this variation into account in a stochastic way in a network setting. Instead of only the type or population mean infectivity and susceptibility, we also regard the distribution of these quantities. We explore effects of variation in infectivity and susceptibility on the probability of a major outbreak occurring.

In this chapter we consider an SIR description of infection spread (see Chapter 1). For notation and terminology we refer to Sections 4.2 and 5.2. We construct a model for the spread of an infection on a (undirected) network/graph $G = (V, E)$, where the individuals/vertices have random infectivity and susceptibility. Individuals that are connected by an edge are neighbours. If G is a complete graph (a graph with edges between every pair of vertices), the model gives a description of the spread of an infection in a randomly mixing population.

The random graphs are constructed as follows: With some abuse of notation we see G as the network with the undirected edges replaced by two directed edges. We assign independently to each vertex/individual, v a pair of random weights $(w_v, \bar{w}_v)$, denoting functions of the infectivity and susceptibility respectively. The weights w_v and $\bar{w}_v$ need not be independent. We denote the (directed) edge from v_1 to v_2 by v_1v_2. The edge v_1v_2 is open with probability $\kappa_G(w_{v_1}, w_{v_2})$, otherwise it is said to be closed. Conditioned on the weights assigned to the individuals, the states (open or closed) of the edges are independent. An open edge v_1v_2 in the random graph corresponds with the event that if v_1 becomes infected, v_1 will have at least one contact with individual v_2, that will make v_2 infectious, if v_2 had not already been infected before. A closed edge v_1v_2 in the random graph implies that v_1 did not have such a contact with v_2. The individuals that are ultimately removed (i.e. the individuals that were infected during the epidemic), are the individuals that can be reached by an open path from the initial infective individuals in the graph representation.

Assume that a susceptible individual v has probability $\bar{w}_v$ of becoming infected every time it has a contact with an infective individual. Furthermore, if v becomes infective it has contacts with every given neighbour at rate w_v^* during a random infectious period of w_v' time units. Then $\kappa_G(w_{v_1}, \bar{w}_{v_2})$ should

be chosen to be $1 - \exp[-w_{v_1}\bar{w}_{v_2}]$, where $w_v = w_v^* w_v'$.

We can also consider an alternative description for infection spread. Assume that there is at most one contact from an infectious individual to a given neighbour during its infectious period, or that only at the first contact of an infective individual with a given neighbour the infection may be transmitted. This last assumption is proposed in some models for the spread of HIV [50, 72], where the number of sexual contacts per couple can be ignored and only the number of partners of individuals is of importance. In those models the probability of an infectious contact from v_1 to v_2 is given by $\kappa_G(w_{v_1}, \bar{w}_{v_2}) = w_{v_1}\bar{w}_{v_2}$, with $w_v = 1 - \exp[w_v^*]$, where w^* is defined as in the previous paragraph. In large randomly mixing populations it is usually assumed that the contact rate of a pair of individuals scales with $\mathcal{N}^{-1}$, where $\mathcal{N}$ is the number of individuals in the population. With this assumption multiple contacts of a pair of individuals are rare, and the difference of the proposals for κ_G made in this and the previous paragraph is of order $\mathcal{N}^{-2}$.

If the probability that $w(v) = \bar{w}(v)$ is 1 for all vertices v, and if $\kappa_G(w_1, w_2)$ is symmetric in w_1 and $\bar{w}_2$, then the probability that there is an open edge from v_1 to v_2 is equal to the probability that there is an open edge from v_2 to v_1. By [25] and Chapter 1 we know that the cluster consisting of vertices that can be reached by an open path from v, is distributed as the vertices in the open cluster containing v of the undirected graph, where the probability of an edge to be open in the undirected graph is equal to the probability that a corresponding edge in the directed graph is open. Furthermore, if G is the complete graph, then the models described above are special cases of the inhomogeneous random graphs, described in [21].

In the next section we construct random graphs of which the random graphs that we use for describing the spread of infections in inhomogeneous populations are special cases. This construction is the main subject of this chapter. It opens doors for further analysis of the spread of infections in inhomogeneous populations and for applying the existing theory on inhomogeneous random graphs [21, 22, 23] to epidemics.

In the third section we give a way to compare probabilities that at least one path in a set of paths is open, under the condition that the probability that edge v_1v_2 is open is factorisable, i.e. $\kappa_G(w_1, \bar{w}_2)$ can be factorised as $\kappa_G(w_1, \bar{w}_2) = \kappa_G^{(1)}(w_1)\kappa_G^{(2)}(\bar{w}_2)$, where $\kappa_G^{(1)}$ and $\kappa_G^{(2)}$ are two arbitrary functions. This can be applied to epidemics, in the sense that infections with a given

expected infectivity and susceptibility can be compared. If the infectivity, $w(v)$ and the susceptibility, $\bar{w}(v)$ are independent of each other, then we prove that the random graphs with fixed w and $\bar{w}$, have the largest expected size of the cluster of vertices that can be reached by open paths from v, and the largest probability that the size of this cluster is of the same order as the number of individuals, i.e. the probability of a major outbreak is maximal. This result is an extension of a result presented in [51], where a similar comparison result is given for infection models, where all individuals have the same susceptibility by assumption.

6.2 Inhomogeneous random graphs: The model

In this section we construct the random graphs that we use to describe the spread of an infection. We also give the notation and terminology needed in the rest of the chapter.

Let $G = (V, E)$ be a non-random directed graph with countable vertex set V and (directed) edge set E. In most applications $|V| = \mathcal{N}$ is large but finite. The vertices are ordered and denoted by $v_1, v_2, \cdots$. Edges are denoted by their start and end-vertex as $v_i v_j$. If the edge $v_i v_j$ exists, then the edge $v_j v_i$ exists as well (i.e. $v_i v_j \in E \Leftrightarrow v_j v_i \in E$). A path of length n in G is a sequence of edges in E of the form $(v_{i_1} v_{i_2}, v_{i_2} v_{i_3}, v_{i_3} v_{i_4}, \cdots, v_{i_n} v_{i_{n+1}})$. A path is self-avoiding if $v_{i_k} \neq v_{i_l}$ for $k \neq l$. A path of length n is a loop, if the truncation of length $n-1$ is a self-avoiding path and $v_{i_1} = v_{i_{n+1}}$.

Let $(W, \bar{W})$ be a positive real random vector with joint distribution function F. The random variables W and $\bar{W}$ need not be independent. Furthermore, let $\mathcal{W} := (\mathcal{W}_i)_{i\geq 1} := ((W_i, \bar{W}_i))_{i\geq 1}$ be an infinite sequence of random vectors, where $(\mathcal{W}_i)_{i\geq 1}$ are independently and identically distributed as $(W, \bar{W})$. Let $\mathbf{w} = ((w_i, \bar{w}_i))_{i\geq 1}$ be a realisation of $\mathcal{W}$. If the number of vertices is larger than i, the realisations w_i and $\bar{w}_i$ are called the weights of vertex v_i and they represent the infectivity and susceptibility of the vertex. If i exceeds the number of vertices, w_i and $\bar{w}_i$ do not have a biological meaning, but it is useful to define $\mathbf{w}$ as an infinite series in order to analyse growing sequences of graphs.

Note that in our terminology every individual has an infectivity assigned to it. It is interpreted as the rate at which other individuals are contacted during its infectious period if the individual ever becomes infected.

We define the kernel κ_G as a function on $\mathbb{R}_+ \times \mathbb{R}_+$, taking values in $[0,1]$. Unless explicitly stated otherwise the function κ_G is continuous. Conditioned on $\mathbf{w}$, the probability that the edge v_iv_j is open is given by $p_{ij} = \kappa_G(w_i, \bar{w}_j)$. If an edge is not open it is closed. Conditioned on $\mathbf{w}$, the states (open or closed) of edges are independent of each other. For finite graphs, $G = (V, E)$, the probability measure $\mathbb{P} := \mathbb{P}_{(G,\kappa_G,F)}$ is the joint measure on the space of states of the edges and weights of the vertices, $\{\text{open,closed}\}^E \times (\mathbb{R}_+ \times \mathbb{R}_+)^V$, where the conditional probability that the edge v_iv_j is open is $p_{ij} = \kappa_G(W_i, \bar{W}_j)$. We use the notation $\mathbb{E}$ for the expectation under the probability measure $\mathbb{P}$.

In general it is not possible to define $\mathbb{P}$ on an infinite complete graph, while keeping it useful for biological applications: Because the marginal probability of an edge to be open is the same for all edges and the degree of every vertex is the same, the number of outgoing open edges of each vertex is either 0 or infinite, which implies that the corresponding infection cannot spread at all or infinitely many individuals will be directly infected by only one initial infective individual. However, it is possible to define $\mathbb{P}$ on countable regular graphs with only finite degrees (see Chapter 5), because this graph does not suffer from the disadvantages of the complete graph.

The expectation $p_G := \mathbb{E}[\kappa_G(W_1, \bar{W}_2)]$, is the marginal probability that a given edge is open. Furthermore, $S_G := (p_G)^{-1}\mathbb{E}[\kappa_G(W_1, \bar{W}_2)\kappa_G(W_2, \bar{W}_3)]$ is the marginal probability that a given self-avoiding path of length 2 is open (i.e. all edges in the path are open), conditioned on the event that the first edge in the path is open.

Remarks

1. Let $|V|$ be finite. If $\mathcal{W}$ can only take values in $[0,1]^2$ and $\kappa_G(w, \bar{w}) = w\bar{w}$, then $S_G = \mathbb{E}(W\bar{W})$. Let A_G be the number of self avoiding paths of length 2 divided by the number of edges. The product A_GS_G is the reproduction ratio R_V of [13], where R_V is defined similar to R_* as defined in Chapter 4. If G is a regular infinite graph where every individual has k outgoing edges (and k ingoing edges because of symmetry), then we can extend the definition to $A_G = k - 1$. (The arguments are similar to the arguments for the definition of ϕ for infinite networks in Section 4.2).

2. If the graph G is a finite complete graph, $\mathcal{N} = |V|$ and $\kappa_G(w_1, \bar{w}_2) = 1 - \exp[-\mathcal{N}^{-1}w_1\bar{w}_2]$, then we obtain a special case of the epidemic models

discussed in the introduction (with w_1 and $\bar{w}_1$ replaced by $\mathcal{N}^{-1/2}w_1$ and $\mathcal{N}^{-1/2}\bar{w}_1$ respectively). This model is comparable to the model of [13]. The cluster of vertices in G that can be reached by open paths from v, is distributed as the cluster of ultimately removed individuals in the model of [13].

3. As stated in the introduction, if $W = \bar{W}$, then the cluster of vertices that can be reached from v by open paths has the same distribution for directed and undirected graphs. If furthermore the graph is complete, we obtain the model of [21]. If $|V| = \mathcal{N}$ and

$$\kappa_{\mathcal{N}}(w_1, w_2) := \kappa_G(w_1, w_2) = w_1 w_2 (\mathcal{N} + w_1 w_2)^{-1},$$

we get the generalised random graph model of [22]. In [23], $p_{ij} = w_i w_j (\sum_{v \in V} w_v)^{-1}$, so p_{ij} depends on the weights of all of the vertices. However, if $\mathcal{N}$ is large and $\mathbb{E}(W) = 1$, then by the law of large numbers, $p_{ij} \approx \kappa_{\mathcal{N}}(w_i, w_j) = \mathcal{N}^{-1} w_i w_j$, is a function of the weights of the end vertices only. Note that this κ can be used as a model for the alternative description of infection spread, given in the introduction.

The pairwise differences of the three proposals for $\kappa_{\mathcal{N}}(w_1, w_2) := \kappa_G(w_1, w_2)$; $\kappa_{\mathcal{N}}(w_1, w_2) = w_1 w_2 (\mathcal{N} + w_1 w_2)^{-1}$, $\kappa_{\mathcal{N}}(w_1, w_2) = 1 - \exp[-\mathcal{N}^{-1} w_1 w_2]$ and $\kappa_{\mathcal{N}}(w_1, w_2) = \mathcal{N}^{-1} w_1 w_2$ are of the order $\mathcal{N}^{-2}$ for $\mathcal{N} \to \infty$.

6.3 A comparison theorem for the random graphs with factorisable $\kappa_G(w, \bar{w})$

In [51] the spread of an infection on a network is studied, where the susceptibility of all individuals is 1 and where the total infectivity is random. Variance in total infectivity may arise by variance in the length of the infectious period. The probability that edge $v_1 v_2$ is open, only depends on a function of the infectious period of v_1. The marginal probability that an edge, with v_1 as its first index is open is said to be p_1. It is proved that for epidemic processes with fixed p_i, the process with no variance in total infectivity is a worst case scenario, in the sense that the probability of a major outbreak as well as the probability that a given individual will become infected in the course of the epidemic is maximal.

In this section we generalise this result from [51] to models on inhomogeneous random graphs with factorisable kernel $\kappa_G(w_1, \bar{w}_2) = \kappa_G^{(1)}(w_1)\kappa_G^{(2)}(\bar{w}_2)$, where $\kappa_G^{(1)}(w_1)$ and $\kappa_G^{(2)}(\bar{w}_2)$ are functions depending on G. The arguments used in this section cannot immediately be generalised to models with arbitrary κ_G. In the rest of this section we set $\kappa_G(w_1, \bar{w}_2) = w_1\bar{w}_2$. We do not lose generality by choosing this kernel, because we can replace the random vector $(W, \bar{W})$ by $(\kappa_G^{(1)}(W), \kappa_G^{(2)}(\bar{W}))$.

The strategy is as follows: We replace the graph G by another graph $\hat{G}$. The states of the edges in $\hat{G}$ only depend on the weight of the first vertex of the edge. We show that the probability of a configuration of states of edges in G is the same as the probability of a corresponding configuration in $\hat{G}$. We can obtain results on $\hat{G}$ by using arguments similar to the ones used in [51].

Let $\tilde{G} = (\tilde{V}, \tilde{E})$ and $\bar{G} = (\bar{V}, \bar{E})$ be copies of G, where vertices $\tilde{v}_i$ and $\bar{v}_i$ correspond to vertex v_i. We consider a new directed graph $\hat{G} = (\hat{V}, \hat{E})$, where $\hat{V} = \tilde{V} \cup \bar{V}$ and $\hat{E} = \tilde{E} \cup \bar{E} \cup \{\tilde{v}\bar{v}; v \in V\}$. The edges from $\tilde{V}$ to $\bar{V}$ are open with probability 1. The edge $\tilde{v}_i\tilde{v}_j$ in $\tilde{G}$ is open with probability w_i. The edge $\bar{v}_i\bar{v}_j$ in $\bar{G}$ is open with probability $\bar{w}_i$. Conditioned on the weights of the vertices, the states of the edges in $\hat{E}$ are independent of each other. The sequence of pairs $\mathbf{w}$ is defined as in the previous section, where W and $\bar{W}$ both take values in $[0, 1]$ with probability 1. Let $\hat{\mathbb{P}}$ be the corresponding joint probability measure on the space of states of the edges and weights of the vertices in $\hat{G}$, $\{\text{open,closed}\}^{\hat{E}} \times (\mathbb{R}_+ \times \mathbb{R}_+)^{\hat{V}}$.

The epidemiological interpretation of the graph $\hat{G}$ is as follows. An open edge from $\tilde{v}_1$ to $\tilde{v}_2$ means that if individual v_1 (in G) becomes infected in the course of the infection (or is an initial infective individual), it will have at least one contact with v_2 during its infectious period. An open edge from $\bar{v}_2$ to $\bar{v}_1$ means that, if there is at least one contact from individual v_1 to individual v_2 (in G) during the infectious period of v_1, the contact is successful and v_2 becomes infective if it has not been infected before.

Let E_1 be a subset of E and $\tilde{E}_1$ (respectively $\bar{E}_1$) be the corresponding subsets of $\tilde{E}$ (respectively $\bar{E}$). Let $\bar{E}_1^* = \{\bar{v}_i\bar{v}_j \in \bar{E}; \bar{v}_j\bar{v}_i \in \bar{E}_1\}$ be the subset of $\bar{E}$ consisting of all edges of $\bar{E}_1$ in the opposite direction. The probability that all of the edges in E_1 are open is equal to the probability that all of the edges in $\tilde{E}_1$ and all of the edges in $\bar{E}_1^*$ are open. A direct consequence of this is that the probability that a given path in G, ξ from v_i to v_j is open is equal to the probability that the path $\hat{\xi}$ in $\hat{G}$ is open, where $\hat{\xi}$ exists of successively

the path in $\tilde{G}$ corresponding with ξ, the edge $\tilde{v}_j\bar{v}_j$ and the path in $\bar{G}$ from $\bar{v}_j$ to $\bar{v}_i$ corresponding to "ξ in the opposite direction".

Let $\Omega := \{\text{open,closed}\}^E$ and $\hat{\Omega} := \{\text{open,closed}\}^{\hat{E}}$. Furthermore, let Ξ be a set of paths in G and define $\hat{\Xi} = \{\hat{\xi}; \xi \in \Xi\}$. The set Ξ_n is the set of truncations at length n of paths in Ξ, and $\hat{\Xi}_n = \{\hat{\xi}; \xi \in \Xi_n\}$, so paths in $\hat{\Xi}_n$ have at most length $2n+1$. Let

$$\mathcal{C}^{\Xi} = \lim_{n\to\infty} \bigcup_{\xi\in\Xi_n} \{\omega \in \Omega; \xi \text{ is open}\} \tag{6.1}$$

and

$$\hat{\mathcal{C}}^{\hat{\Xi}} = \lim_{n\to\infty} \bigcup_{\hat{\xi}\in\hat{\Xi}_n} \{\hat{\omega} \in \hat{\Omega}; \hat{\xi} \text{ is open}\}. \tag{6.2}$$

The set $\mathcal{C}^{\Xi}$ is the set where for every n at least one of the paths in Ξ_n is open.

If W and $\bar{W}$ are independent random variables taking values in $[0,1]$, then we can directly apply the results of [51], because the graph $\hat{G}$ together with the realisations of the states of $\hat{E}$ is a locally dependent random graph. This is interesting in its own right: Let W and $\bar{W}$ be random variables taking values in $[0,1]$ and $\kappa_G(w,\bar{w}) = w\bar{w}$. Let $\mathbf{P}$ be the set of probability measures with the property that W and $\bar{W}$ are independent and $\mathbb{E}(W\bar{W}) = T$, where T is some constant. Let the probability measure $\mathbb{P}^* \in \mathbf{P}$ be the probability measure having the property that $Var(W) = Var(\bar{W}) = 0$. Observe that for every constant $c > 0$ replacing $(W,\bar{W})$ by $(cW, c^{-1}\bar{W})$ does not change the $\mathbb{P}$ measure on the states of the edges. Putting the observations of this paragraph together, brings us at the following theorem (compare this with Theorem 4.1 of [51]):

Theorem 6.3.1 *For every $\mathbb{P} \in \mathbf{P}$ holds that*

$$\mathbb{P}(\mathcal{C}^{\Xi}) \leq \mathbb{P}^*(\mathcal{C}^{\Xi})$$

for any set of paths Ξ in G.

In an epidemiological setting this theorem means that we can compare infections with random infectivity and susceptibility: We compare infections with a given marginal probability for the event that if individual v_1 is infected in the course of the infection, it has at least one contact with individual v_2 during its infectious period that infects v_2 if v_2 was still susceptible at the

moment of the contact. If we consider a set of possible infection routes in the graph G, then the probability that the infection actually spreads following one of those routes is maximal if there is zero variance in susceptibility and infectivity. If v_1 is the single initial infective, then by considering the set of all self-avoiding paths in G with v_1 as the first vertex of the first edge in the path, and v_2 as the second vertex of the last edge in the path, we obtain that the probability that v_2 will become infective during the outbreak is maximal if there is no variance in susceptibility and infectivity of individuals. This implies that R_* (see Chapter 4) is maximal for fixed susceptibility and infectivity. If we consider the set of all self-avoiding paths of length n, with v_1 as the first vertex of the first edge in the paths, then we obtain that the probability that the infection is still spreading after n infection steps is minimal if there is zero variance in susceptibility and infectivity. This implies that the probability of a major outbreak is maximal if the susceptibility and infectivity are fixed.

Define E_v as the edges in G with first end-vertex v. The sets $\hat{E}_{\hat{v}}$ with $\hat{v} \in \hat{V}$ are defined similarly. If W and $\bar{W}$ are dependent, then we cannot use Theorem 6.3.1 and the results of [51] immediately, because the states of edges in $\hat{E}_{\tilde{v}}$ and edges in $\hat{E}_{\bar{v}}$ are not independent, which is required in the construction of locally dependent random graphs. However we can still obtain a result similar to Theorem 2.1 of [51]. Again we need some notation. The zero-function $\hat{z}_v(A,B) = \hat{z}_v(A,B;\hat{\mathbb{P}})$ is a function from all pairs of disjoint, finite, possibly empty subsets of E_v to $[0,1]$ and is defined by

$$\hat{z}_v(A,B) = \hat{\mathbb{P}}\Big(\Big\{\bigcap_{e\in\tilde{A}}\{e \text{ is closed}\}\Big\} \cup \Big\{\bigcap_{e'\in\bar{B}}\{e' \text{ is closed}\}\Big\}\Big),$$

where $\tilde{A} \subseteq \tilde{E}$ and $\bar{B} \subseteq \bar{E}$ are sets of edges that correspond with $A \subseteq E_v$ and $B \subseteq E_v$. The notation $\hat{z}_v^{(1)} \leq \hat{z}_v^{(2)}$ means that $\hat{z}_v^{(1)}(A,B) \leq \hat{z}_v^{(2)}(A,B)$ for all disjoint $A, B \subseteq E_v$. Using this notation we can prove the following theorem:

Theorem 6.3.2 *Let $(\hat{z}_v^{(i)})_{v\in V}$ be the zero functions under the probability measure $\hat{\mathbb{P}}^{(i)}$ defined on the graph G. Let $\hat{\Xi}$ be defined as just above equation (6.1). If $\hat{z}_v^{(1)} \leq \hat{z}_v^{(2)}$ for all v, then for every set of paths Ξ in G it holds that*

$$\hat{\mathbb{P}}^{(2)}(\hat{\mathcal{C}}^{\hat{\Xi}}) \leq \hat{\mathbb{P}}^{(1)}(\hat{\mathcal{C}}^{\hat{\Xi}}),$$

and

$$\mathbb{P}^{(2)}(\mathcal{C}^{\Xi}) \leq \mathbb{P}^{(1)}(\mathcal{C}^{\Xi}).$$

The epidemiological interpretation of this theorem is that the probabilities that at least one of the paths in a set of possible infection paths can be compared and partly ordered for different probability measures, only using the distribution of the infectivities and susceptibilities of the individuals.

Proof. The second inequality in the theorem is a direct consequence of the first inequality and the relation between paths in G and in $\hat{G}$ discussed in this section. The proof of the first part of the theorem is essentially the same as the proof of Theorem 2.1 of [51] and is divided into 4 steps.

(i) The first step is to assume that Ξ is a finite set of finite paths in G and that $\hat{z}_{v_1}^{(1)} \leq \hat{z}_{v_1}^{(2)}$ and $\hat{z}_v^{(1)} = \hat{z}_v^{(2)}$ for all $v \in V \setminus v_1$.

We assume that $|\hat{E}| < \infty$. This is not a restriction, because the state of only a finite number of edges in $\hat{E}$ is relevant for the event $\hat{\mathcal{C}}^{\Xi}$. For $\hat{A} \subseteq \hat{E}$ let $\omega_{\hat{A}}$ be an element in the space $\Omega_{\hat{A}} := \{\text{open,closed}\}^{\hat{A}}$. Let $\hat{E}_{v_1}^c := \hat{E} \setminus \{\hat{E}_{\tilde{v}_1}, \hat{E}_{\bar{v}_1}\}$. In order to keep the formulas readable, we use $\Gamma := \hat{E}_{v_1}^c$. For $i = 1, 2$ it holds that

$$\hat{\mathbb{P}}^{(i)}(\hat{\mathcal{C}}^{\hat{\Xi}}) = \sum_{\omega_\Gamma \in \Omega_\Gamma} \hat{\mathbb{P}}^{(i)}(\hat{\mathcal{C}}^{\hat{\Xi}} | \omega_\Gamma) \hat{\mathbb{P}}^{(i)}(\omega_\Gamma).$$

Note that $\hat{\mathbb{P}}^{(i)}(\omega_\Gamma)$ is the same for $i = 1$ and $i = 2$. The term $\hat{\mathbb{P}}^{(i)}(\hat{\mathcal{C}}^{\hat{\Xi}} | \omega_\Gamma)$ only depends on the states of the edges in $\hat{E}_{\tilde{v}_1}$ and $\hat{E}_{\bar{v}_1}$. It is straightforward to see that if $\hat{z}_{v_1}^{(1)} \leq \hat{z}_{v_1}^{(2)}$, then

$$\hat{\mathbb{P}}^{(1)}(\hat{\mathcal{C}}^{\hat{\Xi}}) - \hat{\mathbb{P}}^{(2)}(\hat{\mathcal{C}}^{\hat{\Xi}}) = \sum_{\omega_\Gamma \in \Omega_\Gamma} \left[\hat{\mathbb{P}}^{(1)}(\hat{\mathcal{C}}^{\hat{\Xi}} | \omega_\Gamma) - \hat{\mathbb{P}}^{(2)}(\hat{\mathcal{C}}^{\hat{\Xi}} | \omega_\Gamma) \right] \hat{\mathbb{P}}^{(1)}(\omega_\Gamma) \geq 0,$$

which proves the theorem for the special case considered.

(ii) In the second step, we still assume that Ξ is a finite set of finite paths in G, but now $\hat{z}_v^{(1)}$ and $\hat{z}_v^{(2)}$ may differ at more than one vertex. It is straightforward to construct an ordered series of probability measures, $(\hat{\mathbb{P}}^{(*;i)})_{1<i<n}$, such that $\hat{\mathbb{P}}^{(*;1)} = \hat{\mathbb{P}}^{(1)}$, $\hat{\mathbb{P}}^{(*;n)} = \hat{\mathbb{P}}^{(2)}$ and two subsequent measures $(\hat{\mathbb{P}}^{(*;i)})$ have a different zero-function at only one vertex. We can apply the (i) part of this proof to two subsequent probability measures, which proves the theorem for the special case considered in this step.

(iii) If $\hat{\Xi} = \{\hat{\xi}_1, \hat{\xi}_2, \cdots\}$ is an infinite countable set of finite paths in $\hat{E}$, then we define for $k \in \mathbb{N}$ the truncated sets $\hat{\Xi}^k = \{\hat{\xi}_1, \hat{\xi}_2, \cdots \hat{\xi}_k\}$. For $k \to \infty$, $\hat{\mathcal{C}}^{\hat{\Xi}^k} \to \hat{\mathcal{C}}^{\hat{\Xi}}$, therefore $\hat{\mathbb{P}}^{(2)}(\hat{\mathcal{C}}^{\hat{\Xi}}) \leq \hat{\mathbb{P}}^{(1)}(\hat{\mathcal{C}}^{\hat{\Xi}})$.

(iv) For general Ξ we use the definition

$$\hat{\mathcal{C}}^{\hat{\Xi}} = \lim_{n\to\infty} \bigcup_{\hat{\xi}\in\hat{\Xi}_n} \{\hat{\omega} \in \hat{\Omega}; \hat{\xi} \text{ is open}\}. \tag{6.3}$$

the paths in $\hat{\Xi}_n$ are all finite. Using the previous steps in the proof, we obtain that

$$\hat{\mathbb{P}}^{(2)}(\hat{\mathcal{C}}^{\hat{\Xi}}) \leq \hat{\mathbb{P}}^{(1)}(\hat{\mathcal{C}}^{\hat{\Xi}}).$$

This completes the proof of the theorem.

Remark: If for some finite n, $\hat{\Xi}_n$ is an infinite set of disjoint paths, then the statement in the theorem is not very informative, because if $\hat{z}_v^{(i)}$ is the constant function 1 for no $v \in V$, then $\hat{\mathbb{P}}^{(i)}(\hat{\mathcal{C}}^{\hat{\Xi}_n}) = 1$ for $i = 1, 2$. However for many applications $\hat{\Xi}_n$ is finite for all n.

One important observation is that we cannot use this theorem directly as is done in [51] to proof that for all probability measures on $\{\text{open,closed}\}^E \times (\mathbb{R}_+ \times \mathbb{R}_+)^V$, with $\mathbb{E}(W) = T_1$ and $\mathbb{E}(\bar{W}) = T_2$, where T_1 and T_2 are fixed, the case where $W = T_1$ and $\bar{W} = T_2$ gives a worst case scenario. That is, if $\mathbf{P}$ is the relevant set of probability measures and $\mathbb{P}^* \in \mathbf{P}$ is the probability measure having the property that $Var(W) = Var(\bar{W}) = 0$, then the result of Theorem 6.3.1 does not necessarily hold.

A basic counterexample is given by considering one path of length 2, ξ and two measures $\mathbb{P}^{(1)}(W = \bar{W} = T) = 1$ and $\mathbb{P}^{(2)}(W = \bar{W} = 1) = T$, $\mathbb{P}^{(2)}(W = \bar{W} = 0) = 1 - T$, where $0 < T < 1$. We have

$$\begin{aligned} \mathbb{P}^{(1)}(\xi \text{ is open}) &= T^4, \\ \mathbb{P}^{(2)}(\xi \text{ is open}) &= T^3. \end{aligned}$$

However, if $\mathbb{E}(\bar{W}|W = w)$ is decreasing in w, then the result of Theorem 6.3.1 holds. To prove this we use that if Z is a random variable and g_1 and g_2 are positive and decreasing functions then

$$\mathbb{E}[g_1(Z)g_2(Z)] \geq \mathbb{E}[g_1(Z)]\mathbb{E}[g_2(Z)]$$

(see e.g. Lemma 4.2 of [51] or page 184 of [47]). By repeated use of this inequality we obtain

$$
\begin{aligned}
\hat{z}_v(A,B) &:= \hat{\mathbb{P}}\Big(\Big\{\bigcap_{e\in\tilde{A}}\{e \text{ is closed}\}\Big\}\cup\Big\{\bigcap_{e'\in\bar{B}}\{e' \text{ is closed}\}\Big\}\Big)\\
&= \hat{\mathbb{E}}\Big(1-[1-(1-W)^{|A|}][1-(1-\bar{W})^{|B|}]\Big)\\
&= \hat{\mathbb{E}}\Big(1-[1-(1-W)^{|A|}][1-\hat{\mathbb{E}}((1-\bar{W})^{|B|}|W)]\Big)\\
&\geq \hat{\mathbb{E}}\Big(1-[1-(1-W)^{|A|}][1-(1-\hat{\mathbb{E}}(\bar{W}|W))^{|B|}]\Big)\\
&= \hat{\mathbb{E}}\Big((1-\hat{\mathbb{E}}(\bar{W}|W))^{|B|}\Big)+\hat{\mathbb{E}}\Big((1-W)^{|A|}[1-(1-\hat{\mathbb{E}}(\bar{W}|W))^{|B|}]\Big)\\
&\geq \hat{\mathbb{E}}\Big((1-\hat{\mathbb{E}}(\bar{W}|W))^{|B|}\Big)+[1-\hat{\mathbb{E}}(W)]^{|A|}\hat{\mathbb{E}}\Big(1-(1-\hat{\mathbb{E}}(\bar{W}|W))^{|B|}\Big)\\
&= [1-\hat{\mathbb{E}}(W)]^{|A|}+(1-[1-\hat{\mathbb{E}}(W)]^{|A|})\hat{\mathbb{E}}\Big((1-\hat{\mathbb{E}}(\bar{W}|W))^{|B|}\Big)\\
&\geq [1-\hat{\mathbb{E}}(W)]^{|A|}+(1-[1-\hat{\mathbb{E}}(W)]^{|A|})[1-\hat{\mathbb{E}}(\bar{W})]^{|B|}\\
&= 1-\Big(1-[1-\hat{\mathbb{E}}(W)]^{|A|}\Big)\Big(1-[1-\hat{\mathbb{E}}(\bar{W})]^{|B|}\Big)\\
&= \hat{\mathbb{P}}^*\Big(\Big\{\bigcap_{e\in\tilde{A}}\{e \text{ is closed}\}\Big\}\cup\Big\{\bigcap_{e'\in\bar{B}}\{e' \text{ is closed}\}\Big\}\Big).
\end{aligned}
$$

By applying Theorem 6.3.2 we obtain the desired result.

On complete graphs the expectation $S_G = \mathbb{E}(W\bar{W})$ is determining the reproduction ratio R_V of [13]. We therefore expect that S_G is the relevant quantity and that we can generalise Theorem 6.3.1 if we compare graphs with $S_G = T$ for some fixed T. This brings us to the following conjecture.

Conjecture 6.3.3 *Let $\mathcal{P}$ be the set of probability measures with the property that $\mathbb{E}(W\bar{W}) = T$. Let $\mathbb{P}^\dagger \in \mathcal{P}$ be the probability measure having the property that $Var(W) = Var(\bar{W}) = 0$. Let Ξ be a set of finite self-avoiding paths with first vertex of the first edge v_i and the second vertex of the last edge is v_j. Individual v has weights $(w_v, \bar{w}_v)$ assigned to it. Then for all $\mathbb{P} \in \mathcal{P}$ and for any graph G, it holds that*

$$\mathbb{P}(\mathcal{C}^\Xi|w_i,\bar{w}_j) \leq \mathbb{P}^\dagger(\mathcal{C}^\Xi|w_i,\bar{w}_j).$$

If Ξ is a set of infinite self-avoiding paths with first vertex of the first edge v_i. Then for all $\mathbb{P} \in \mathcal{P}$ and for any infinite graph G, it holds that

$$\mathbb{P}(\mathcal{C}^\Xi|w_i) \leq \mathbb{P}^\dagger(\mathcal{C}^\Xi|w_i).$$

This conditioning on the weights of the first and the last vertex is because $\mathbb{E}(W\bar{W})$ is the probability that a given contact of an individual with an infectious neighbour causes the individual to become infective, and that after that the individual has contacts with another given individual. However, the initial infective in a path need not be infected by a neighbour, and it is not important whether the last individual in a path spreads the infection to other individuals or not. Note that this conditioning is of no importance for the results on R_* (Chapters 4 and 5), because only the number of secondary cases infected by a secondary infective individual is important. If the presence of infinite paths has positive probability taking the conditioning into account, then the presence of infinite paths also has positive probability if one does not use the conditioning. It is possible to replace the conditioning on the actual weights of the "end-vertices of the path" by keeping the expected values of the separate weights (infectivity and susceptibility) of those vertices fixed for the models that we compare.

6.4 A short discussion on the model for general $\kappa_G(w_1, \bar{w}_2)$ and possible further research

In the previous section we constructed a graph $\hat{G} = (\hat{V}, \hat{E})$, where the states of the edges in $\hat{E}_{\hat{v}}$ depend only on the weight of $\hat{v}$. We cannot construct such a graph if there does not exist two functions $\kappa_G^{(1)}$ and $\kappa_G^{(2)}$, such that $\kappa_G(w, \bar{w}) = \kappa_G^{(1)}(w)\kappa_G^{(2)}(\bar{w})$.

An interesting question is, whether a variant of Theorem 6.3.1 can be formulated in this general context? E.g. with the following definitions from the introduction,

$$\begin{aligned} p_G &:= \mathbb{E}(\kappa_G(W_1, \bar{W}_2)) \\ S_G &:= (p_G)^{-1}\mathbb{E}(\kappa_G(W_1, \bar{W}_2)\kappa_G(W_2, \bar{W}_3)), \end{aligned}$$

let $\mathcal{P}$ be defined as the set of probability measures with $S_G = s$, where s is a given constant. Let $\mathbb{P}^\dagger \in \mathcal{P}$ be the probability measure having the property that $Var(W) = Var(\bar{W}) = 0$. Let Ξ be a set of finite self-avoiding paths with first vertex of the first edge v_i and the second vertex of the last edge is v_j. Is

it still true that for all $\mathbb{P} \in \mathcal{P}$ and for any graph G, it holds that

$$\mathbb{P}(\mathcal{C}^{\Xi}|w_i, \bar{w}_j) \leq \mathbb{P}^{\dagger}(\mathcal{C}^{\Xi}|w_i, \bar{w}_j)?$$

Furthermore, if Ξ is a set of infinite self-avoiding paths with first vertex of the first edge v_i. Is it still true that for all $\mathbb{P} \in \mathcal{P}$ and for any infinite graph G, it holds that

$$\mathbb{P}(\mathcal{C}^{\Xi}|w_i) \leq \mathbb{P}^{\dagger}(\mathcal{C}^{\Xi}|w_i)?$$

In this chapter the existing theory on inhomogeneous random graphs has not yet been used. This is mainly because this existing theory is only about complete graphs G, while the results of this chapter are interesting on general graphs G. However, it is interesting whether and how the methods to obtain results in inhomogeneous random graphs discussed in [21, 22, 23] can be generalised to the inhomogeneous random graphs discussed in this chapter. One might expect that interesting results, like the extinction probability and the basic reproduction ratio, R_0 and even branching process approximations can be obtained for infection spread in inhomogeneous, randomly mixing populations. Further research on this subject may give new insights in mathematical epidemiology and is therefore desirable.

Bibliography

[1] Albert, R. and Barabasi, A.L. (2002), Statistical mechanics of complex networks, *Reviews of Modern Physics* **74** 47-96.

[2] Andersson, H. (1999), Epidemic models and social networks, *The Mathematical Scientist* **24(2)** 128-147.

[3] Andersson, H. and Britton, T. (2000), Stochastic epidemic models and their statistical analysis, *Springer Lecture Notes in Statistics* **151**, New York: Springer Verlag.

[4] Ball, F.G. (1986), A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models, *Advances in Applied Probabiltiy* **18** 289-310.

[5] Ball, F.G. and Donnelly, P. (1995), Strong approximations for epidemic models, *Stochastic Processes and their Applications* **55** 1-21.

[6] Ball, F.G.; Mollison, D. and Scalia-Tomba, G. (1997), Epidemics with two levels of mixing, *Annals of Applied Probabiltiy* **7** 46-89.

[7] Ball, F.G. and O'Neill, P.D. (1999), The distribution of general final state random variables for stochastic epidemic models, *Journal of Applied Probability* **36(2)** 473-491.

[8] Ball, F.G. and Britton, T. (2005), An epidemic model with exposure-dependent severities, *Journal of Applied Probability* **42(4)** 932-949.

[9] Ball, F.G. private communication.

[10] Bauch, C.T. (2005), The spread of infectious diseases in spatially structured populations: An invasory pair approximation, *Mathematical Biosciences* **198** 217-237.

[11] Bearman, P.S.; Moody, J. and Stovel, C. (2004), Chains of Affection: The structure of adolescent romantic and sexual networks, *American Journal of Sociology* **110(1)** 44-91.

[12] Becker, N.G. and Hasofer, A.M. (1997), Estimation in epidemics with incomplete observations, *Journal of the Royal Statistical Society, B(Methodological)* **59(2)** 415-429.

[13] Becker, N.G. and Starczak, D.N. (1998), The effect of random vaccine response on the vaccination coverage required to prevent epidemics, *Mathematical Biosciences* **154** 117-135.

[14] Becker, N.G. and Utev, S. (2002), Protective vaccine efficacy when vaccine response is random, *Biometrical Journal* **44(1)** 29-42.

[15] Benjamini, I. and Berger, N. (2001), The diameter of long-range percolation clusters on finite cycles, *Random Structures and Algorithms* **19** 102-111.

[16] Benjamini, I.; Kesten, H.; Peres, Y. and Schramm, O. (2004), Geometry of the uniform spanning forest: Transitions in dimensions $4, 8, 12, \ldots$, *Annals of Mathematics* **160(2)** 465-491.

[17] Berger, N.; Borgs, C.; Chayes, J. and Saberi, A. (2005), On the spread of viruses on the internet, *In Proceedings of the 16-th ACM-SIAM Symposium on Discrete Algorithms, 2005*

[18] Biskup, M (2004), On the scaling of the chemical distance in long-range percolation models, *Annals of Probability* **32(4)** 2938-2977.

[19] Biskup, M (2004), Graph diameter in long-range percolation, preprint at arXiv:math.PR/0406379

[20] Bollobás, B. (2001), *Random Graphs second edition*, Cambridge University Press.

[21] Bollobás, B.; Janson, S. and Riordan, O (2005) The phase transition in inhomogeneous random graphs, preprint at arXiv:math.PR/0504589

[22] Britton, T.; Deijfen, M. and Martin-Löf, A. (2006), Generating simple random graphs with prescribed degree distribution, *Research reports, Stockholm University: Mathematical Statistics* 2005:2.

[23] Chung, F. and Lu, L. (2002), Connected components in random graphs with given expected degree sequences, *Annals of Combinatorics* **6(2)** 125-145.

[24] Coppersmith, D.; Gamarnik, D. and Sviridenko, M. (2002), The diameter of a long-range percolation graph, *Random Structures and Algorithms* **21** 1-13.

[25] Cox, J.T. and Durrett, R. (1988), Limit theorems for the spread of epidemics and forrest fires, *Stochastic Processes and their Applications* **30(2)** 171-191.

[26] Daley, D.J. and Gani, J. (1999), *Epidemic Modelling: An Introduction*, Cambridge University Press.

[27] Diekmann, O. and Heesterbeek, J.A.P. (2000), *Mathematical Epidemiology of Infectious Diseases*, Chichester: John Wiley & Son.

[28] Durrett, R. (2006), *Random Graph Dynamics*, To be published by Cambridge University Press, manuscript available at `http://www.math.cornell.edu/~durrett/RGD/RGD.html`.

[29] Ferguson, N.M.; Donnelly C.A. and Anderson R.M. (2001), The Foot-and-Mouth epidemic in Great Britain: Pattern of spread and impact of interventions, *Science* **292** 1155-1160.

[30] Ferguson, N.M.; Donnelly C.A. and Anderson R.M. (2001), Transmission intensity and impact of control policies on the Foot and Mouth epidemic in Great-Britain, *Nature* **413** 542-547.

[31] Ganesh, A.; Massoulie, L. and Towsley, D. (2005), The effect of network topology on the spread of epidemics, *In Proceedings of IEEE INFOCOM, 2005*

[32] Grimmett, G.R. and Stirzaker, D.R. (1992), *Probability and Random Processes second edition*, New York: Oxford University Press Inc.

[33] Grimmett, G.R. (1999), *Percolation second edition*, Berlin, Heidelberg, New York: Springer Verlag.

[34] Guttorp, P. (1991) *Statistical Inference for Branching Processes*, New York: John Wiley & Sons.

[35] Hall, P. and Heyde, C.C. (1980) *Martingale Limit Theory and its Applications*. New York: Academic Press.

[36] van der Hofstad, R.; Hooghiemstra, G. and van Mieghem, P. (2005), Distances in random graphs with finite variance degrees, *Random Structures and Algorithms* **26** 76-123.

[37] Jackson, M.O. and Yariv, L (2005), Diffusion on social networks, *Economie Publique* **16**.

[38] Jacob, C. and Peccoud, J. (1998) Estimation of the parameters of a branching process from migrating binomial observations, *Advances in Applied Probability* **30(4)** 948-967.

[39] Jagers, P. (1975), *Branching Processes with Biological Applications*, London: John Wiley & Sons.

[40] Kao, R.R. (2002), The role of mathematical modelling in the control of the 2001 FMD Epidemic in the UK, *Trends in Microbiology* **10 (6)** 279-286.

[41] Keeling, M.J. (1999), The effects of local spatial structure on epidemiological invasions, *Proceedings of the Royal Society, London B* **266** 859-867.

[42] Keeling, M.J.; Woolhouse, M.E.J.; Shaw, D.J.; Matthews, L.; Chase-Topping, M.; Haydon, D.T.; Cornell, S.J.; Kappey, J.; Wilesmith, J. and Grenfell B.T. (2001), Dynamics of the 2001 UK Foot and Mouth Epidemic - Dispersal in a heterogeneous landscape, *Science* **294** 813-817.

[43] Keeling, M.J. and Eames, K.T.D. (2005), Review. Networks and epidemic models , *Journal of the Royal Society Interface* **2(4)** 295-307.

[44] Keeling, M.J. (2005), The implications of network structure for epidemic dynamics, *Theoretical Population Biology* **67(1)** 1-8.

[45] Kesten, H. (1980), The critical probability of bond percolation on the square lattice equals 1/2, *Communications in Mathematical Physics* **74(1)** 41-59.

[46] Kermack, W.O. and McKendrick, A.G. (1927), Contributions to the mathematical theory of epidemics, *Bulletin of Mathematical Biology* **53 (1/2)** 33-118

[47] Kingman, J.F.C. (1978), Uses of exchangibility, *Annals of Probability* **6** 183-197.

[48] Klinkenberg, D. (2003) Mathematical Epidemiology and the Control of Classical Swine Fever Virus, *PhD Thesis Utrecht University*

[49] Klinkenberg, D.; Everts-van der Wind, A.; Graat, E.A.M. and De Jong, M.C.M. (2003), Quantification of the effect of control strategies on classical swine fever epidemics, *Mathematical Biosciences* **186(2)** 145-173.

[50] Knolle H. (2004), A discrete branching process model for the spread of HIV via steady sexual partnerships, *Journal of Mathematical Biolgy* **48** 423-443.

[51] Kuulasmaa, K. (1982), The spatial general epidemic and locally dependent random graphs, *Journal of Applied Probability* **19** 745-758.

[52] van Lieshout, M.N.M. (2000), *Markov Point Processes and Their Applications*, London: Imperial College Press.

[53] Liggett, T.M. (1999), *Stochastic interacting systems: Contact, voter and exclusion processes*, Berlin: Springer Verlag.

[54] Liljeros, F.; Edling, C.R.; Amaral, L.A.N.; Stanley, H.E. and Aberg, Y. (2001), The web of human sexual contacts, *Nature* **411** 907-908.

[55] Meester, R. and Roy, R. (1996), *Continuum Percolation*, Cambridge University Press.

[56] Meester, R.; de Koning, J.; de Jong, M.C.M. and Diekmann, O. (2002), Modelling and prediction of classical swine fever epidemics, *Biometrics* **41** 178-184.

[57] Meester, R. and Trapman, J.P. (2005), Estimation in Branching Processes with restricted observations, *submitted:* available at `http://www.math.vu.nl/~rmeester/preprints/bpschatting.pdf`

[58] Mollison, D. (1977), Spatial contact models for ecological and epidemic spread, *Journal of the Royal Statistical Society, Series B* **39(3)** 283-326.

[59] Mollison, D. (1991),Dependence of epidemic and population velocities on basic parameters , *Mathematical Biosciences* **107** 255-287.

[60] Murray, J.D.; Stanley, E.A. and Brown, D.L. (1986), On the spatial spread of rabies among foxes, *Proceedings of the Royal Society, London B* **229** 111-150.

[61] Murray, J.D. (1989), *Mathematical Biology*, Berlin: Springer Verlag.

[62] Murrell, D.J.; Dieckmann, U. and Law, R. (2004), On moment closures for population dynamics in continuous space, *Journal of Theoretical Biology* **229(3)** 421-432.

[63] Newman, M.E.J. (2002), Spread of epidemic disease on networks, *Physical Review E* **66(1)** 016128.

[64] Newman, M.E.J. (2003), Ego-centered networks and the ripple effect, *Social Networks* **25** 83-95.

[65] Ormerod, P. and Roach, A.P. (2004), The Medieval inquisition: Scale-free networks and the suppression of heresy, *Physica A* **339(3-4)** 645-652.

[66] Petermann, T. and De Los Rios, P. (2004), Cluster approximations for epidemic processes: a systematic description of correlations beyond the pair level, *Journal of Theoretical Biology* **229(1)** 1-11.

[67] Rand, D.A. (1999), Correlation equations and pair approximations for spatial ecologies. In *Advanced Ecological Theory* (ed. J. McGlade), pp. 99-143, London: Blackwell Scientific Publishing.

[68] Sellke, T. (1983), On the Asymptotic Distribution of the Size of a Stochastic Epidemic, *Journal of Applied Probability* **20(2)** 390-394.

[69] Stegeman, A.; Elbers, A.R.W.; Smak, J. and de Jong, M.C.M. (1999), Quantification of the Transmission of Classical Swine Fever Virus between Herds during the 1997-1998 Epidemic in The Netherlands, *Preventive Veterinary Medicine* **42** 219-234.

[70] Trapman, P.; Meester, R. and Heesterbeek, J.A.P. (2004) A branching model for the spread of infectious animal diseases in varying environments, *Journal of Mathematical Biology* **49** 553-576.

[71] Trapman, P. Work in progress.

[72] Watts, C.H. and May, R.M. (1992), The influence of concurent partnerships on the dynamics of HIV/AIDS, *Mathematical Biosciences* **108** 89-104.

[73] Watts, D.J. and Strogatz, S.H. (2002), Collective dynamics of 'small-world' networks, *Nature* **393** 440-442.

Samenvatting

Over stochastische modellen voor de verspreiding van infecties

De werkelijke dynamiek van de verspreiding van een besmettelijke ziekte is meestal te complex om te beschrijven in een wiskundig analyseerbaar model. Om die reden zijn vereenvoudigingen nodig die een kwalitatief beeld geven van de verspreiding en die als gereedschap kunnen dienen voor het analyseren van computerintensieve simulaties van deze verspreiding.

Het meest gebruikte stochastische (kanstheoretische) model voor de verspreiding van besmettelijke ziekten is het stochastische SIR (Vatbaar (S) $\rightarrow$ Infectieus (I) $\rightarrow$ Immuun (R))-model. In dit model veronderstellen we dat ieder individu zich in één van deze drie "toestanden" (S, I of R) bevindt. Als een vatbaar individu een contact heeft met een infectieus individu, wordt het zelf ook infectieus met een bepaalde kans. Een infectieus individu herstelt na een exponentieel verdeelde infectieuze periode. Contacten tussen ieder tweetal individuen vinden met een zelfde intensiteit plaats (de aanname van een willekeurig mengende populatie). Een hersteld individu is immuun en blijft dat voor altijd.

Dit proefschrift handelt over uitbreidingen van dit standaardmodel. Voor deze uitbreidingen wordt veelvuldig gebruik gemaakt van bekende kansprocessen, zoals vertakkingsprocessen.

Vertakkingsprocessen zijn ontwikkeld om stambomen te modelleren. Men wilde weten hoe groot de kans is dat een gegeven achternaam uitsterft en hoe lang dit gemiddeld duurt. Er wordt aangenomen dat de aantallen zonen van alle mannen onafhankelijk en identiek verdeeld zijn. Ook de leeftijden waarop mannen hun zonen krijgen zijn onafhankelijk en identiek verdeeld.

Een beginnende epidemie in een grote willekeurig mengende populatie kan

ook met een vertakkingsproces beschreven worden. In een grote populatie is de kans dat er in een beginnende epidemie al contacten worden gemaakt tussen infectieuze en niet-vatbare individuen erg klein. Als we deze kans helemaal negeren dan wordt het verloop van het aantal infectieuze individuen in een epidemie exact beschreven met een vertakkingsproces, waarbij de individuen die door een infectieus individu besmet zijn, gezien worden als zijn kinderen.

In Hoofdstuk 2 beschouwen we de verspreiding van klassieke varkenspest (KVP) tussen bedrijven. We beschouwen hier de bedrijven als de individuen. De verspreiding van de infectie tussen bedrijven hangt af van de verspreiding binnen de bedrijven. Als er meer dieren op een bedrijf besmet zijn, neemt de kans dat de besmetting wordt overgebracht naar een ander bedrijf toe. Ook neemt de kans toe dat de besmetting ontdekt wordt en het bedrijf geïsoleerd en geruimd wordt. Daarom modelleren we de verspreiding van KVP op twee niveaus: binnen de bedrijven en tussen de bedrijven.

We zijn geïnteresseerd in de verspreiding tussen de bedrijven. Echter, door de spreiding binnen de bedrijven is de lengte van de infectieuze periode niet meer exponentieel verdeeld. Verder is de intensiteit waarmee contacten die tot een besmetting van andere bedrijven leiden niet meer constant, maar hangt af van de tijd sinds de eerste besmetting op het bedrijf. Daarom is voor de verspreiding tussen bedrijven een uitbreiding van het standaard SIR-model gebruikt.

Omdat klassieke varkenspest grote gevolgen heeft voor het dierenwelzijn en de economie, worden er maatregelen genomen die de verspreiding van de infectie moeten afremmen. Deze maatregelen zorgen voor extra uitdagingen in het modelleren, want de intensiteit waarmee contacten plaatsvinden verandert door deze maatregelen tijdens de epidemie. Dit zorgt ervoor dat we het model nog verder moeten aanpassen om tot een realistische beschrijving van de verspreiding te komen. In Hoofdstuk 2 ontwerpen we het model dat de uitbreidingen uit de vorige en deze alinea bevat. Met dit model kunnen we kwalitatieve voorspellingen doen over het effect van maatregelen die genomen worden om de verspreiding te remmen.

In het model van Hoofdstuk 2 gebruiken we verschillende parameters. In werkelijkheid weten we de waarden van deze parameters vaak niet op het moment dat een epidemie uitbreekt en moeten we deze schatten tijdens de epidemie zelf. In Hoofdstuk 3 gaan we in op de vraag of en hoe we deze parameters kunnen schatten met behulp van de beperkte informatie die doorgaans

beschikbaar is als een epidemie nog aan het spreiden is. Uit de bestaande literatuur over vertakkingsprocessen is bekend dat we, als we alleen maar weten hoeveel individuen op ieder moment infectieus zijn, maximaal twee parameters kunnen schatten. Maar zelfs dit aantal infectieuze individuen is meestal niet bekend. We weten wel het aantal individuen waarvan we waargenomen hebben dat ze besmet zijn en die op het moment van deze observatie stoppen met infectieus zijn, door isolatie of ruiming. In dit hoofdstuk tonen we aan dat we, verrassend genoeg, met deze observaties uiteindelijk drie parameters kunnen schatten.

In de hoofstukken 4, 5 en 6 laten we de aanname van willekeurige menging los. In plaats daarvan beschouwen we de verspreiding van een besmettelijke ziekte over een sociaal netwerk. Voor deze netwerken nemen we aan dat er alleen contacten mogelijk zijn tussen individuen die ook in het sociale netwerk met elkaar verbonden zijn.

In Hoofdstuk 4 wordt de in de wiskundige epidemiologie veel gebruikte methode van "paarbenaderingen" geanalyseerd. Deze methode negeert kansprocessen en is erg grof, maar het geeft wel een beeld van het verloop van de verspreiding in de tijd. We tonen aan dat andere methoden welkom en misschien zelfs noodzakelijk zijn voor verdere analyse van de epidemie.

Eén van de andere manieren om de verspreiding over een netwerk te modelleren is de volgende. We vervangen het netwerk, waarover de infectie zich verspreidt, in het model door een nieuw netwerk. Dit nieuwe netwerk heeft bepaalde eigenschappen gemeen met het oorspronkelijke netwerk, zoals de verdeling van het aantal verbindingen per individu, maar is verder willekeurig. Het nieuwe netwerk biedt de gelegenheid om de verspreiding van de infectie te analyseren. We kunnen met deze methode rekening houden met de kansprocessen die een rol spelen bij de verspreiding. Het is zodoende mogelijk om te bepalen hoe groot de kans op een kleine uitbraak (een epidemie waarbij niet meer dan enkele individuen worden besmet) is en hoe groot het verwachte aantal besmette individuen is tijdens de gehele epidemie.

In de bestaande netwerk-modellen is de enige eigenschap van het sociale netwerk, die wordt overgenomen in het nieuwe netwerk, de verdeling van het aantal verbindingen per individu. Door de individuen te vervangen door huishoudens van willekeurige omvang, wordt in Hoofdstuk 5 ook de mogelijkheid gegeven om het aantal driehoeken in het sociale netwerk over te nemen in het nieuwe netwerk. Driehoeken zijn belangrijk want die geven in feite weer

dat de vrienden van mijn vrienden vaak ook mijn vrienden zijn. Modellen voor de verspreiding van een besmettelijke ziekte op dit nieuwe netwerk zijn goed te analyseren en ook op dit netwerk is het mogelijk om de kans op een kleine uitbraak te bepalen.

In het laatste hoofdstuk van dit proefschrift worden contacten tussen infectieuze en vatbare individuen verder geanalyseerd. Totnogtoe is in deze inleiding aangenomen dat elk contact met een vaste kans leidt tot een besmetting. Echter, in de praktijk zijn sommige individuen meer vatbaar voor een infectie dan anderen. Ook scheiden sommige individuen meer infectieus materiaal uit dan anderen en hebben dus een hogere infectiviteit. We modelleren dit als volgt: ieder individu krijgt twee getallen toegekend. De eerste geeft de vatbaarheid weer, de tweede de infectiviteit. De kans dat bij een contact tussen een vatbaar en een infectieus individu daadwerkelijk een besmetting plaatsvindt wordt gegeven door een functie die afhangt van twee variabelen, namelijk de infectiviteit van het infectieuze individu en de vatbaarheid van het vatbare individu. De vectoren (vatbaarheid, infectiviteit) voor de verschillende individuen zijn identiek en onafhankelijk verdeeld, maar de vatbaarheid en infectiviteit van één individu kunnen wel gecorreleerd zijn.

Een variant van het bovenstaande epidemiologische model is in de wiskundige literatuur al bestudeerd in de context van willekeurige netwerken. Het is echter nooit eerder in verband gebracht met epidemieën. Nu dit verband wel is aangetoond, kunnen resultaten voor de willekeurige netwerken toegepast worden op modellen voor de verspreiding van besmettelijke ziekten.

Acknowledgements

Writing acknowledgments is a tough job. I have many people to thank for their scientific and non-scientific support. In fact, they are too many and omissions in this part of the thesis can hardly be avoided, therefore I do not even try to do so, but I will only mention the most important contributors.

First of all I want to thank my promotors, Ronald Meester and Hans Heesterbeek. Ronald, thanks for being a patient teacher both in the research part and in the scientific writing part. I think it is mainly because of your enthusiasm that I rediscovered the beauty of mathematics instead of only its usefulness for applications. I would never have learned to appreciate the theory of interacting particle systems without your help. Apart from the working-contacts I have enjoyed the debates/discussions/talks on the interaction of science and theology. Hans, you were always there to discuss ideas and to stimulate me to develop and to elaborate my brainwaves. I am very glad that you gave me the freedom in choosing the topics of my research. I enjoyed our discussions and the social contacts have always been very pleasant to me.

The members of the reading committee, Frank Ball, Odo Diekmann, Klaas van Harn and Mart de Jong, have given very useful comments on the manuscript of this thesis, their comments made it possible to correct and to improve the text and have triggered enough ideas for a couple of years of further research.

Without the help of my colleagues I would not have been able to finish this thesis. Space does not permit to thank everyone that deserves it, therefore I reserve my praises to the closest contacts. Alexis Gillett is a nice roommate: we have had a multitude of discussions, even some on mathematics. His beautiful simulation applets have helped me to get more insight in spatial stochastic systems. I thank Geurt Jongbloed for the collaboration in the lectures and exams and for the social talks on church-matters. My "train partner" Corrie Quant has always been good company and has contributed to this thesis by

commenting on my ideas. Misja Nuijens has helped me a lot with his bright insights and his ability to create a pleasant atmosphere. Of my Utrecht colleagues Don Klinkenberg deserves the most credits. Our discussions were very useful for the research in this thesis and Don's knowledge and help on Mathematica matters were essential for the progress of my research. I also thank my other roommates, Birgitte Ask and Liesbeth van der Waaij. Michiel van Boven's comments on draft versions of papers and parts of this thesis have definitely been of great worth.

I am indebted to all the people in the reading groups I joined. I want to give a special place to Stan Alink for the talks and email chats we had and of course for his wonderful way of commenting on the state of the world.

For making the layout of this thesis I have used the advise (and the master files) of Rachel Brouwer, Misja Nuyens and Wouter Kager.

Personal words are there for the people that showed me that there is more than a scientific life. Wouter en Ingrid, I am very thankful for both of you being friends indeed. It is good to have you as a housemate and a "housemate-in-law". Bea, it has always been a pleasure to have dinner with you. I am glad that you want to be a "paranimf". Pa, Ma, Mirjam en Nelly, you have given me a real home from birth (well, not all of you) to now. Pa, thank you for being a paranimf. I look forward to see you in tailcoat.

Finally, the late Rich Mullins once sang: "I believe, what I believe is what makes me what I am. I did not make it, no it is making me" This statement may also be part of my creed. Therefore, I want to end this thesis as I started it, with words from the Good Book.

The LORD is my shepherd; I shall not want.
He maketh me to lie down in green pastures: he leadeth me beside the still waters.
He restoreth my soul: he leadeth me in the paths of righteousness for his name's sake.
Yea, though I walk through the valley of the shadow of death, I will fear no evil: for thou art with me; thy rod and thy staff they comfort me.
Thou preparest a table before me in the presence of mine enemies: thou anointest my head with oil; my cup runneth over.
Surely goodness and mercy shall follow me all the days of my life: and I will dwell in the house of the LORD for ever.
(Psalm 23, KJV)

Pieter